DOCUMENT RESUME

ED 059 823 RC 005 991

AUTHOR	Taylor, Benjamin J.; Helmkamp, John
TITLE	Arizona Indian Demographic Data: Needs and Recommendations.
SPONS AGENCY	Department of Housing and Urban Development, Washington, D.C.
PUB DATE	[71]
NOTE	105p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	*American Indians; *Demography; Human Resources; Needs; Population Distribution; *Research Methodology; *Reservations (Indian); *Sampling

ABSTRACT

Included in this report on Arizona Indian demographic data are "an evaluation of several recent studies of Indian populations" and "an extensive analysis of methods for obtaining and maintaining accurate data in the future." Recommended methods by which accurate population data for the smaller reservations should be maintained are included in the first portion of the report, while 4 sampling techniques (simple random sampling, stratified simple random sampling, single-stage cluster sampling, and systematic sampling) for use on the larger reservations are evaluated and detailed in the second part of this paper. The appendices provide a sample questionnaire for Indian demographic studies, a delineation of notation and formulas used for sample survey designs, and simulated sampling results. (LS)

ED 059823

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ARIZONA INDIAN DEMOGRAPHIC DATA:

NEEDS AND RECOMMENDATIONS

RECEIVED
MAR 30 1972
NMSU
E.R.I.C.

Prepared By

Benjamin J. Taylor
and
John Helmkamp

For

Planning Division

Arizona Department of Economic
Planning and Development

[1971]

The Preparation of This Report was Financed in part through a Comprehensive Planning Grant from the United States Department of Housing and Urban Development

RC 005991

FOREWORD

Information is vital to successful program design and implementation. This study began under an assumption that reasonably reliable data on Arizona's Indian populations could be readily obtained or developed. The research effort demonstrated that securing and maintaining accurate demographic data was both more difficult and more important, because of the wide discrepancies found in available estimates, than originally envisioned. Because of the magnitude and significance of the problem, this report has resulted in a program design rather than just a report on Indian populations.

The Department of Economic Planning and Development is pleased to be able to publish this report and expresses its gratitude to the authors for the work done. Dr. Benjamin J. Taylor is Chairman of the Economics Department and Dr. John Helmkamp is Chairman of the Accounting Department in the College of Business at Arizona State University.

This report is directed toward those agencies of government with programs and responsibilities for aiding the Indian nations in Arizona. It is requested that these agencies examine this report and consider implementation of a recommended program for maintaining current and reliable demographic data.

Robert G. Worden
Executive Director
Arizona Department of
Economic Planning and
Development

TABLE OF CONTENTS

		Page
FOREWORD		ii
LIST OF TABLES		iv
CHAPTER 1	INTRODUCTION	1
CHAPTER 2	OVERALL EVALUATION OF INDIAN DEMOGRAPHIC DATA FOR ARIZONA	5
CHAPTER 3	GENERAL SAMPLING STRATEGY	29

3

LIST OF TABLES

Table Number		Page
1	Comparison of Population Figures Reported By Four Principal Agencies	8
2	Planned and Actual Sample Sizes	14
3	Actual Standard Errors for a 20 Percent Proportion By Reservation	16
4	Number of Communities on Each Reservation	17
5	An Example of a Sample Selection Table	19
6	Assumed Costs With Simple Random Sampling	42
7	Assumed Correlation Between Vehicle Registration and Indian Income	46
8	Strata Established for 1000 Simulated Individual Indian Incomes	50
9	Sample Size Per Stratum	51
10	Assumed Costs With Stratified Simple Random Sampling	52
11	Assumed Costs With Single Stage Cluster Sampling	60
12	Assumed Costs With Systematic Sampling	67

4

CHAPTER 1

INTRODUCTION

The State of Arizona has one of the largest populations of Reservation Indians in any state in the nation. The 1970 Census of the Population reports 95,812 Indians residing in Arizona. At April 1, 1960, the Census report 83,387 Indians in Arizona. The great majority of these Indians live on reservations, at least most of the time.

The 95,812 Indians residing in Arizona at April 1, 1970, represented 5.4 percent of the state's total population of 1,772,489. This ratio is reduced from the 6.4 percent of the total Arizona population represented by Indians in 1960, but remains a significant portion of the population. More importantly, Indian residents of Arizona form a much larger share of the state's low income residents. Indians suffer disproportionately from unemployment and underemployment, poor housing, problems in the delivery of health services and educational opportunities.

The many problems faced by Arizona's Indian populations are recognized, but a lack of accurate demographic data frequently thwarts efforts of state and federal agencies to help solve them, and it obscures the actual magnitude of various opportunities as well as deficiencies.

The State of Arizona is interested in aiding the Indian Tribes in their development efforts. However, the ability of the state to assist in plans and programs for improving the economic and social environment of the reservations is heavily dependent upon accurate data. The magnitude of the information problem can be illustrated by comparing recent estimates with U. S. Census figures.

Current Data Problems

In a recently completed working paper for the Department of Economic Planning and Development, Gonzales and Johnson made an analysis of Reservation population "census" figures being used at the time by the Bureau of Indian Affairs.[1/] The Tribal Census total population figures being used were 107,849 in June of 1969 and 104,018 in March of 1970. Other sources cited on page 11 of that paper indicated 1970 estimates of 102,000 to 104,000. When these estimates are compared with the U. S. Census count of 95,812, it is obvious better data are needed, particularly when it is remembered that the U. S. Census count includes all Indians in the state, not just those on the reservations.

Apache County in Arizona has one of the larger Indian populations, and an examination of the problem of maintaining accurate estimates in this county is revealing. The U. S. Census reported 32,298 residents in Apache County at April 1, 1970, and of this total, 23,994 or 74.3 percent, were Indians. In 1960, the Census count for the county was

---

1/ Val Gonzales and Eugene Johnson, Integration of the Indian Reservations with Arizona's Economic Planning and Development Process, Arizona Department of Economic Planning and Development, (Phoenix, 1970), pp. 8-10.

30,438 residents, of which 22,814, or 75 percent, were Indian. The Unemployment Compensation Division of the Arizona Employment Security Commission uses accepted Bureau of the Census methods in preparing population estimates for Arizona's counties. For July 1, 1969, it estimated the population of Apache County as 46,500, substantially above the 1970 Census count of 32,298.

Information Needs

The reason for the apparent over estimate of the population of Apache County reflected above is not difficult to explain. Records of vital statistics show the ten-year natural increase (number of births over number of deaths) in the County's population was 11,897. Also, of the 14,644 births in the County between 1960 and 1970, 11,487 or 78.4 percent were Indians. Therefore, if the U. S. Census counts are accurate, there was a substantial migration of Indians from the county. An analysis of Census counts and vital statistics data indicates a total out-migration from the county during the decade of 10,037 persons.

Without a method for maintaining information on these population movements, it is difficult to assist in the development of the reservations. Prospective industries require reliable labor force data before investing in plants and equipment. These data are also needed for establishing priorities for public investments in programs and facilities to alleviate

the problems of manpower training, housing, health and welfare which beset the Indian peoples.

Objectives

This report includes, first, an evaluation of several recent studies of Indian populations and, secondly, an extensive analysis of methods for obtaining and maintaining accurate data in the future. Recommendations of methods by which accurate population data for the smaller reservations should be maintained are included in the first portion of the report. Various sampling techniques for use on the larger reservations are evaluated and methods are explained in detail in the second portion of this paper. It is hoped that resources will be available to maintain accurate data on the Indian populations to enable state and federal agencies to assist in solving some of the pressing problems faced by reservation Indians.

CHAPTER 2

OVERALL EVALUATION OF INDIAN DEMOGRAPHIC DATA FOR ARIZONA

Indian demographic studies are available, but in limited quantity and quality. Many studies have been conducted on various reservations, but most of them are not generally accessible for review or use. It is clear from the studies, such as are available, that there are several outstanding weaknesses in all of the attempts to develop reliable data. One serious problem is the poor quality of reservation population rolls from which reliable samples may be drawn. Not a single reservation in the State of Arizona maintains a well-edited population roll which is up-to-date at all times. The fact that Indians are far more mobile than is usually realized makes such a task one of monumental proportions. This very basic deficiency makes it extremely difficult to develop data over any prolonged period of time. It should be realized, however, that maintenance of up-dated population rolls on all reservations at all times would be expensive and may be unnecessary.

Previous Studies

Most studies are usually one-shot efforts that more likely than not are only reasonably characteristic of the on-reservation population

during the finite period in which the data were collected. There is no valid reason, for example, to assume that the data collected in June are characteristic of the on-reservation population in any other month of the year. This problem exists because approximately one-half of all Indians reside off their reservations at any given period of time. There is evidence that this one-half of the total population changes substantially from one period of time to another. Thus, longitudinal studies of the demographic characteristics of Indians are desirable if more adequate data are to be developed which will generate more conclusive findings regarding Indian manpower characteristics.

Efforts to use various censuses of several reservations have been made; however, such efforts have proven less reliable than samples because of the vast mobility of Indians. A census does not permit calculation of error whereas sample errors are calculable. Again, a census at one period of time on a particular reservation, assuming all on-reservation Indians are interviewed, could result in substantially different findings from one census to the next because of Indian migration. Relatively inexpensive sampling methods should be utilized to obtain the basic data needed about Indians which would, at the same time, not make subsequent samples prohibitive in the event that a check is desired on the initial sample.

The single most pressing problem that should be resolved before any serious attention can be devoted to developing and maintaining current data on reservation Indians is maintenance of carefully edited population rolls on a selective basis. The tribes themselves are in the best position to accomplish this task. A clerk will probably have to be maintained on the larger reservations who can devote a proportion of her total work time to this function. On small reservations, less time would be required of an employee to fulfill this function. The Department of Economic Planning and Development may find it necessary to assist tribes or some group, such as Indian Development Districts of Arizona, to obtain funds for the purpose. Continuously well-maintained population rolls would permit development of demographic data at minimal cost.

Calculated Cost of Maintaining Rolls

Only eight of the twenty reservations listed on Table 1 would require almost continuous attention to maintain reasonably correct population rolls. All of the others should be developed initially and then edited only once per year, unless more frequent editing is required for sampling. It is assumed that continuous effort would be required to maintain rolls on reservations with populations of 1,500 or more. After the initial rolls are developed for reservations with less than 1,500 population, no more than $500 should be required to bring them

Table 1

Comparison of Population Figures Reported
By Four Principal Agencies

Tribe	IDDA[8] Mar 69	ACIA[9] May 70	BIA[10] Mar 70	IHS[11] Jul 70
Ak Chin	220	240	240	400
Camp Verde	607	314	314	686
Cocopah	99	63	64	-
Colorado River	1,620	1,297	1,297	1,430
Fort Apache	5,712	5,953	5,953	5,800
Fort McDowell	261	280	280	330
Gila River	6,140	5,241	5,241	6,450
Gila Bend	103	-	244	262
Havasupai	365	270	270	-
Hopi	4,966	6,000	4,996	7,250
Hualapai	682	1,028	682	1,520
Kaibab-Paiute	140	60	60	-
Navajo	69,631	71,396	71,396	-
Papago	4,544	5,506	4,688	7,900
Salt River	2,212	2,040	2,040	2,265
San Carlos	4,583	4,404	4,404	5,050
Yavapai-Prescott	85	90	90	90
San Xavier	558	-	574	-
Totals[12]	102,528	104,182[c]	102,803	39,833[d]

[c]Total not complete--several figures not available.

[d]Total not complete--several figures not available.

[8]Four Corners Regional Commission Technical Assistance Project, Arizona, Preliminary State Development Plan, Arizona Department of Economic Planning and Development, March 1969, Phoenix.

[9]Arizona Commission of Indian Affairs, Tribal Directory, 1970, a pamphlet.

[10]Bureau of Indian Affairs, Reservation Programs, March 1970, op. cit.

[11]Indian Health Service, Sacaton Service Unit, Program Plan, Phoenix Area and Personal Interview with Phil Pepper and Charles Green.

[12]Although population figures are not available for all the tribes in the Arizona Commission for Indian Affairs and the Indian Health Service columns, the discrepancies reflected in each agency's figures for the respective tribe are apparently obvious.

Source: Taken from Val Gonzales and Eugene Johnson, "Integration of the Indian Reservations with Arizona's Economic Planning and Development". Mimeographed Paper, August 1970.

completely up-to-date. Very little expense would be involved for reservations such as Ak Chin, Cocopah and Yavapai-Prescott. Knowledgeable on-reservation informants could edit small lists in about one day. It is estimated that approximately $7,500 would be required to obtain initial working lists for all the Arizona reservations, exclusive of the Navajo. This figure does not include the full time cost of one staff person's salary for approximately three months.

Editing difficulties will accelerate with reservation size. Even knowledgeable informants will be unable to maintain a perfectly accurate list. Keeping up with births and deaths is not an impossible task, but there is a severe limitation in attempting to learn which Indians are on or off the reservation at any moment of time. Reservations are usually composed of vast acreages which make it impossible for even the more informed Indians to know of all of the migration. But consistent effort can minimize the number of mistakes. This observation suggests that editing should take place at the time a survey is to be conducted and that continuous editing would be unnecessary. This is particularly the situation if primary emphasis is placed on the age 16 and over population as opposed to the total population of each reservation.

There is still another very important problem associated with attempts to collect Indian reservation demographic data. Indian tribes are very reluctant to cooperate with researchers. Usually, data are

collected within two to three weeks after all other arrangements have been made, but the preliminary arrangements may take as long as six months. Tribal councils require long periods of time prior to deciding to cooperate with some group that wishes to collect data. The Indian is suspicious of the objectives of most groups that want to survey their reservations. Any agency or firm that requires elementary information concerning Indians should be aware of the public relations time that may be required to obtain tribal cooperation. For this reason, it is highly probable that all Indian tribes will not support every request to run a sample. This problem may make it virtually impossible to conduct longitudinal studies to obtain more accurate demographic data. A single study of any given Indian reservation may be the very best that may be expected. It is suggested, therefore, that once permission is granted to sample a given reservation, an attempt should be made to obtain information on as many aspects of Indian life that may possibly be needed for planning. The incremental cost of doing so would also be very small compared to the expense associated with a new sample each time information is needed on any given subject.

EVALUATION OF THREE RESERVATION POPULATION STUDIES

Indian Manpower Resources in the Southwest: A Pilot Study[2/]

Sampling Plan. The overall strategy of the sample involved several steps. First, a questionnaire was developed and tested.

2/ Benjamin J. Taylor and Dennis J. O'Connor, Indian Manpower Resources in the Southwest: A Pilot Study, Arizona State University, (Tempe, Arizona, 1969).

The questions generally concentrated on individual experiences that would permit extensive description of the nature of Indian manpower utilization. Second, population lists were developed for each reservation for the purpose of sample selection. Finally, interviewers from each reservation were selected and trained to deal with the perceivable problems that might arise in the conduct of interviewing. Continuous direction was provided the Indian interviewers during the interview period.

Population Lists: The population lists on the different reservations came from different sources and presented different problems. The most reliable tribal rolls for the two reservations in New Mexico were maintained by the Bureau of Indian Affairs. These tribal rolls provided the primary sampling frame for the Pueblo Laguna and the Pueblo of Acoma. The rolls were first edited by tribal officials and then later by different members of the tribes. Persons not known to the individuals editing the population lists were included in the sampling frames.

The inclusion of unknowns in the sampling frames did not introduce a sampling bias since if they were not on the reservation, even though their name appeared in the sample, the non-response was from a person not in the population of interest. The standard error did tend to rise because of this feature, but remained within manageable limits.

The same procedure was followed on all five reservations included in the study. The basic problem of this approach is that there is no assurance that the population remains the same, even for periods as short as a week or a month. Demographic characteristics, therefore, are subject to violent change.

Sample Size. The sample size for each reservation is related to population characteristics, such as proportions and means, that were to be developed in the survey. Since the questionnaire focuses primarily on frequency of response to certain questions, the sample size was determined with reference to the standard error of the proportion.

The specification adopted for the planned error is that the standard error for a 20 percent characteristic not exceed 1.5 percentage points. This means that a sample estimate of 20 percent should be capable of establishing a 95 percent confidence limit of $20 \pm 1.96\ (1.5)$, or approximately 17.0 - 23.0 percent, for the true value for the entire reservation. This is a relative error of 1.5/20 or 7.5 percent. The relative errors for characteristics exceeding 20 percent will then be less than 7.5 percent. For characteristics occurring with less than 20 percent frequency, the expected sampling errors will exceed the 7.5 percent figure.

Given the restriction that the relative error for 20 percent characteristics not exceed 7.5 percent, random sampling theory is used

to find the requisite sample size. Taking account of the finite sizes of the populations sampled, the error formula is

$$V^2 = \frac{q}{np} \quad \frac{N - n}{N}$$

Where V is the error limitation of .075, i.e., 7.5 percent
p is .20, i.e., 20 percent
q is 1-.20, i.e., 80 percent
N is the size of the population
n is the sample size

Restated to solve for sample size, this formula becomes

$$n = \frac{Nq}{V^2Np + q} = \frac{.80N}{(.075)^2 (.20)N + .80}$$

For some illustrative population sizes, the required sample sizes are shown below:

Population Frame	Sample Size
1000	420
2000	530
3000	580
4000	610
5000	630
Infinite	710

These sample sizes will yield estimates for the mean of a normally distributed variable that have high assurance (95 percent) of being correct within 2.6 percent; 95 percent confidence limits can be established for the mean of a highly skewed variable within 5.4 percent.

The relative errors for percentage characteristics or for totals based upon them will depend on the sample percentage obtained. These sample sizes will provide estimates within 7.6 percent of the true value (with 95 percent confidence) for 50 percent characteristics, within 15 percent of the true value for 20 percent characteristics (the original error specification), and within 22 percent of the true value for 10 percent characteristics.

Table II shows both the sample sizes that were planned and those that actually resulted. Those planned for each reservation were for the purpose of keeping the standard error of the proportion at the level of .015. The planned sample sizes ranged from 500 for the Acoma Reservation, which was the smallest community in the study, to 580 for the Papago Reservation, which was the largest reservation in terms of the 16 years of age and over on-reservation population. Despite the fact that there were more than twice as many names on the Papago population list as compared to the Laguna population list, the difference in the planned sample sizes was only 80.

Table 2

Planned and Actual Sample Sizes

	Planned	Actual
Laguna	527	261
Acoma	500	240
Pagago	580	382
Fort Apache	550	554
San Carlos	524	350

The actual size of the sample for each reservation was affected by non-responses on each reservation. The fact that some of the sampling frames used included many people who no longer resided on the reservation resulted in reduced sample sizes. On the Laguna Reservation, for example, 527 names were selected from the modified population list provided by the Bureau of Indian Affairs. Approximately one-half of those selected were actually on the reservation and as a result the sample size turned out to be 261.

On the Fort Apache Reservation, a somewhat different procedure was followed. Since it was surveyed after both the Laguna and the Acoma Reservations, experience made it seem obvious that the actual sample size would be smaller than the planned sample size unless the sample was inflated. A pretest of the population list was run to determine its accuracy. A sample of 686 was drawn in the hopes of getting 550 responses. This technique worked well since 554 usable responses were obtained from the effort. This is the only reservation for which the standard errors are equal to the desired levels. In all other cases, the standard errors of the proportions are larger than intended. These standard errors are not so large as to make interpretation of the statistical results difficult. The standard errors of the proportions for a 20 percent characteristic, given the actual sample sizes used on each reservation, are reported in Table 3. The preceding discussion

Table 3

Actual Standard Errors for a 20 Percent Proportion
By Reservation
(Planned error was .015)

Laguna	.021
Acoma	.022
Papago	.019
Fort Apache	.015
San Carlos	.020

applies to sampling errors associated with estimates of the population as a whole. For purposes of analysis, various subgroups of the population were examined. Some of these subgroups are relatively small and the sampling errors associated with these subgroups are larger than those of the general population. Information for some of the small subgroups is presented even though sampling errors are large. This is done because some of the information is very suggestive and because little other information on the particular subject is available. The reader is, however, advised to interpret subgroup data with caution. Differences in subgroup data must be large if the differences are to be statistically significant.

Procedure for Drawing Sampling Units

The procedure for selecting sampling units was structured so that the advantages of geographic stratification would be present. The sampling lists were organized so that all people in a given community

were grouped together. The sampling list for each reservation may be viewed as a series of lists representing the smaller communities on the reservation. Each of the smaller communities represents a strata within the list. The number of strata for each reservation is reported in Table 4.

Table 4

Number of Communities on Each Reservation

Laguna	7
Acoma	2
Papago	54
Fort Apache	10
San Carlos	11

The communities vary widely in size. In some cases, the population of a community may be less than 25 persons. In other cases, the population for a community may be 1,000 or more.

The sampling takes advantage of any tendency for the characteristics under investigation to vary systematically by community. These gains in accuracy from effective stratification by community tend to reduce the sampling error below that of an unrestricted random sample of comparable size. On the other hand, inability to completely enumerate names selected in the sample and errors remaining in the edited reservation lists tend to increase sampling error.

Interpenetrating replicate subsamples were drawn from the stratified sampling list. That is, instead of drawing one large sample, a number of smaller samples were drawn. The number of subsamples was set at ten. For example, if a sample of 400 was desired, 10 subsamples of 40 were drawn. The number of subsamples selected is large enough to make possible calculation efficiencies, but it is not too large as to interfere with the stratification of the sampling list.

The first step in the selection of the sampling units was to divide the sampling list into equal zones. The number of zones in the sampling list is equal to the number of elements in a subsample. Each and every zone provides one observation for each of the ten subsamples. Thus, if the sampling list contained 4, 000 names and a sample of 400 is desired, then the sampling list would be divided into 40 zones,each with 100 names.

The selection of the elements to be included in each subsample was made on a systematic basis. The first unit of each subsample was selected at random from the first zone of the sampling list. The subsequent units included in each subsample were selected by adding a number equal to the zoning interval to the random start. An example of a sample selection table is illustrated in Table 5. This table assumes that there is a list of 4, 000 names from which 10 subsamples of 40 each will be drawn. This means that there are 40 zones with 100 names in each zone. Selection of the starting numbers is without replacement.

Table 5

An Example of a Sample Selection Table

Zone	Subsample 1	2	3	4	5	6	7	8	9	10
#0001 - 0100	55	27	68	etc.						
#0101 - 0200	155	127	168	etc.						
#0201 - 0300	255	227	268	etc.						

#3901 - 4000	3955	3927	3968	etc.						

Strength of the Method

There are several strengths of the method as designed. It achieves the advantages of geographic stratification. The larger communities on each reservation are weighted more heavily in terms of sampling units selected than the smaller communities. In addition, by concentrating the interviewers in different communities, any interviewer bias that might be present is spread over a number of subsamples. The method is also convenient for making the calculation of standard errors routine and simple. Since there are hundreds of estimates of proportions for each reservation, this last advantage is of major significance. The standard error of any proportion can be estimated through the examination of subsample proportions. The standard error can be calculated by subtracting the lowest subsample proportion, the highest subsample proportion and dividing by the number of subsamples. The estimate of

$$p = \frac{P\ max - P\ min}{10}$$

the total proportion is computed through the combination of all ten sub-sample frequencies. If the population roll is highly accurate, reasonable estimates may be obtained because the response to various questions will be larger than if the roll from which the sample is drawn does not reflect high accuracy.

The book Indian Manpower Resources in the Southwest: A Pilot Study contains detailed characteristics of reservation populations, details about the employed and unemployed, extent of education and training, consumption and expenditure patterns, skill and industry experiences, and sources of income.

Specific categories include:

1. Population by age group, six, and education attainments.
2. Occupation and industry experiences of the working age populations over a five-year period.
3. Labor force participation rates by sex, age, and educational level.
4. Under utilization of the manpower resource.
5. Labor force status by sex, age and education.
6. Source of training for usual job.
7. Place (on or off reservation) where work is performed.
8. Ability to communicate in the English language.

9. Disabilities preventing individuals from working.

10. Source and amount of income received by individuals and families during the previous year.

11. Consumption and expenditure patterns.

Data collected are useful for a variety of purposes. They will give, if data are current, prospective employers insights into the skill availability of reservation Indians. Also, they will assist employers and relevant government agencies in determining which skills training programs are more likely to be successful.

The data will not provide prospective employers with information regarding the skills possessed by off reservation Indians who would probably return to their reservations if job opportunities should develop. Thus, the book should be taken in certain cases as revealing minimal characteristics of reservation Indian potential for development.

INTEGRATION OF THE INDIAN RESERVATIONS WITH ARIZONA'S ECONOMIC PLANNING AND DEVELOPMENT PROCESS

This is a working paper prepared for the Department of Economic Planning and Development by Val Gonzales and Eugene Johnson. It is comprised primarily of an attempt to find and list existing documents that deal with the demographic features of two reservations - The Papago and The Gila River. Other reservations were included, but the

major effort was devoted to the two reservations mentioned. Much of the effort was devoted to explaining to tribal leaders the need and importance of economic planning and development of reservation economies.

The report provides population estimates of the various Arizona reservations made by four groups. The discrepancies observed in the four estimates may reflect not only the differences in procedures used, but also the different times of the year that each group made the estimate. Nevertheless, the discrepancies reveal the problem facing researchers who attempt to establish sampling devices to obtain valid demographic data.

Much of the report is an account of the difficulties that Gonzales and Johnson encountered in their attempts to obtain the cooperation of tribal governments to develop accurate tribal rolls. The basic reason that widespread discrepancies exist is that the tribes do not maintain adequate rolls themselves and display an unwillingness to cooperate with groups that would welcome the opportunity to do so.

The report is primarily an "in-house" working paper with limited value to others. The document's greatest value is the documentation of the extreme difficulty that researchers can expect to encounter when they try to obtain Indian cooperation in almost any project imaginable. Even Indians who seek tribal assistance find they have considerable

difficulty with the various tribes. Not all tribes, however, exhibit the same degree of difficulty as can be expected from the Papago. The Papago Reservation has had a survey run almost continuously during recent years because of the activities of the Indian Health Service and the University of Arizona. Other tribes have had little of such activity and can be expected to be more cooperative.

NAVAJO MANPOWER SURVEY

This report was a result of work by the Navajo Manpower Survey Task Force sponsored by Navajo Tribe, Indian Health Service, Bureau of Indian Affairs, Arizona State Employment Service, and Office of Navajo Economic Opportunity.[3/]

The Navajo Task Force obtained a population register from the Bureau of Indian Affairs, Navajo Area Office. The population register is maintained by the 72 B. I. A. school districts. The school districts were combined and 18 universes were formed to establish districts of at least 3, 000 persons each (14 years of age and older). Except in the more industrialized sections, at least a 10 percent sample was drawn from the districts.

There was no discussion of difficulties with the drawn sample. That is, no indication was given regarding the quality of the B. I. A. maintained population rolls. School district rolls may not focus on

---

3/ Bernd Schwartz, Navajo Manpower Survey, Arizona State Employment Service (Phoenix, Arizona, January 1970).

families without school age children. Many questions come to mind regarding the quality of the population lists which could lead to substantial bias in the data.

It was indicated that more than a 10 percent sample was drawn in the more industrialized sections of the reservation. No mention was made of adjustments to insure that such a practice did not bias the overall characteristics of the Navajo population. Considerably more information concerning the method used to obtain data is necessary before the study can be adequately judged.

The Navajo Manpower Survey provides a brief description of the evolvement of the socio-economic situation on the Reservation. It also speculates regarding the political structure of the Reservation which has generated considerable unity of the various factions with regard to utilization of their resources. Substantial emphasis is placed on the Navajo desire to attract industry to the reservation as opposed to training of individuals to go to distant areas to compete effectively for jobs with non-Indians. In this regard, the Navajo Manpower Survey proceeds with the same premise as other studies that focus on Indian manpower. The Indian is not different from other groups in the population in that all prefer to remain in more familiar surroundings. Once the basic premise is stated, the study proceeds to describe the characteristics of the Navajo population, aged 14 and over.

Characteristics of the manpower resource focused primarily on the non-student population because of the difficulty found in attempts to contact students. The failure to include students in the study could result in a substantial understatement of the population that might prefer employment, if such were available. That is, many students were enrolled in off-reservation boarding schools who may otherwise have remained on the reservation to work and forego schooling, they were given that choice. However, this situation has been communicated to the reader and does not seriously distort the findings because the student population is treated separately.

The study is similar to Indian Manpower Resources in the Southwest in that it contains data on the population by age grouping, educational attainment, use of the English language, and the desire for training.

The Navajo study also treats the characteristics of the labor force in a manner that is not consistent with either definition used by the United States Department of Labor or the Indian Manpower Resource Study. The Navajo study included a person in the labor force if he indicated he desired to work. This definition overstates the size of the labor force relative to the official definition of that concept. However, this is a debilitating problem only if there is a desire to compare the Navajo to the United States or some subgroup of the country. If the data are to be used only to illustrate what one might reasonably expect to find

on the Navajo Reservation, then the definition is acceptable. Navajo Manpower Survey recognizes this problem.

Labor force participation by age and sex is illustrated in the study. It also provides information on educational attainment of the labor force and calculates median schooling which are useful measures for prospective employers seeking location sites. Navajo occupational experience is included in the study. Such data are highly useful because they indicate the degree of Indian experience with various categories of labor force discipline. Stability of the Indian in an employment relationship can be inferred from the study. Also, the size of the "available pool" of workers is indicated by the data. This can be inferred from the responses of those participating in the survey, but, of course, such responses should be approached with caution. Those who stated they desired employment may, in fact, be unemployable in most occupational categories, if not all of them. Other problems are associated with such responses.

Characteristics of the employed and unemployed Navajos are also provided in convenient manner. These data are broken into the categories of usual occupation, educational attainment, and age and sex.

In brief, the Navajo Manpower Survey provides a reasonable illustration of the reservation Indian's labor force characteristics at the time the survey was made. There were no income and consumption

data, such as were included in the Indian Manpower Resources Study, but the Navajo study had an objective that was somewhat different from it. The Navajo survey is an important document that will serve as a primary source on labor force data until adequate funds become available to replace it. The major problem with the study is the horrendous problems associated with the sample technique. The data should be used with caution.

Bureau of Indian Affairs and Indian Health Service Studies, various Bureau of Indian Affairs offices and Indian Health Service units conduct studies that deal with demographic features of reservation Indians. Most of these tend to be informal working papers with a limited purpose. Some eventually find their way into print, but most become lost in storage. The sampling methods used vary drastically from highly scientific to the naive. Usually, it is almost impossible to obtain information about the method used because of personnel turnover in those agencies, and, thus, most of these documents must be ignored even if they could be uncovered.

The Indian Health Service servicing the Papago Reservation is most active in assembling data for the Papago Indian and should be utilized prior to any surveying anticipated for that tribe. Much of the information needed may be in their possession. The Bureau of Ethnic Studies at the University of Arizona has also assembled data on the

Papago Indian and is also a valuable source of certain types of demographic data which may preclude the necessity of having to sample for certain information.

The other reservations have not had the same magnitude of activity, and, therefore, greater effort will be associated with obtaining basic demographic data. It should be noted that the Papago and Navajo reservations pose the greatest sampling problems. All others are relatively simpler. This is due partly to the relatively larger geographic size of these two reservations and partly to the greater dispersion of their on-reservation populations.

CHAPTER 3

GENERAL SAMPLING STRATEGY

The objective of the sample design stage of the Indian study is to insure that the relevant statistical measures listed on the questionnaire and discussed earlier are estimated on a probilistic basis that insures objective and unbiased results. A wide selection of sample survey designs are available for collecting the important statistics concerning Indian life. Four specific plans are recommended in this section to achieve the survey objectives described in other parts of this study. The methods are: (1) simple random sampling; (2) stratified simple random sampling; (3) cluster sampling; and (4) systematic sampling. Each plan has certain advantages and disadvantages which must be evaluated carefully to insure that an efficient sample design is achieved. An efficient sample plan is one which achieves a high degree of precision per unit of cost. Precision in this sense represents a measure of how close an estimate is to the true value being estimated. The two major dimensions which must be evaluated for this objective are: (1) The amount of variation and skewness in the frequency of occurrence of the statistical measures of interest; (2) The amount of geographical dispersion on a given Indian reservation population. Since the entire

population is not enumerated, the sampler must consider the risk of a sampling error as measured by the standard deviation of the sample mean (hereafter referred to as the standard error). This error arises because we cannot expect the estimated results to be exactly the same as the true values involved so the concern is to control any differences within probilistic limits.

Since the precision of the results obtained from a sample survey depends primarily on the sample size and method of sample selection, different efficiency levels can be achieved by choosing among the sampling plans described in this section. Consequently, the sample size required to attain a desired precision-cost objective can be altered through the choice of the sampling method. When future surveys of Indian populations are made, the sampler should consider the prevailing conditions of a particular project (such as the resources available, objectives of the sample, precision demands, geographical dispersion on the reservation, and the amount of variation between the statistical lives of different Indians) and match them with the characteristics discussed in this section to attain a satisfactory efficiency level. As a result, different plans may be chosen at different times to control the standard error within acceptable boundaries at a reasonable cost, thus achieving varying precision per unit of cost outcomes. Once a given plan is selected as being the most appropriate for a specific reservation,

it normally would be used consistently over time to insure comparable results unless significant changes occur.

A complete sample program is presented in this section for each of the plans to provide a procedural description with which future samples of Indian data can be drawn. The population elements are defined as all Indians who are sixteen years of age or older and the statistics of interest for this study are those listed on the questionnaire shown in Appendix A.

Many of the statistics listed on the questionnaire are either estimates of attributes (frequency of occurrence of some characteristic such as a given income level) or estimates of variables measurements (average or total values of measures of interest, such as the average individual Indian income for a given reservation). Since we did not have an actual situation to work with in this study, certain assumptions were made to provide an illustrative base for describing the various sampling techniques. In a real life application of the sample plans, the questionnaire shown in Appendix A will serve as the instrument with which the vital Indian statistics will be retrieved. However, in the description of the different plans in this section, the focus is on income data exclusively and the hypothetical individual Indian income population shown in Appendix C is used for illustrative purposes. Since most of the statistical measures of interest on the questionnaire shown

in Appendix A are either directly or indirectly related to Indian income, the income focus provides an effective base for administering the survey anyway. This point will be discussed more completely later. The population of 1,000 numbers representing the income levels of the Indians of a hypothetical reservation provides the basis for demonstrating the sample selection process with the various plans. In actual applications, the same analysis would be extended to the other measures recorded on the questionnaire. The income values were simulated from the results of an earlier study by Taylor and O'Connell, Indian Manpower Resources In the Southwest.

We assume that sampling without replacement is used throughout the analysis. Therefore, once a given Indian has been selected, he cannot be chosen again for inclusion in the sample. We also assume that appropriate record keeping of the population rolls will be achieved for each Indian reservation of interest (see an earlier section for a related discussion). Consequently, movements to and from the reservation will be recorded continuously and the population roll will be accurate any time a sample is drawn. Obviously, this factor is extremely important since the sample results of any survey cannot be any more representative of the true population measures of interest than the data recorded on the list from which the sample is drawn. Thus, the only type of error considered in this study is the sampling error. We

recommend strongly that an adequate investment in a record keeping system for the maintenance of reliable population rolls be made and that periodic checks be performed to confirm that the records are correct. If it is discovered that a significant amount of movement on and off the reservation occurs but cannot be effectively recorded, an attempt should be made to determine the impact of this movement on the estimated statistics of interest. One suggestion to accomplish this objective would be to estimate several mean values of measures of interest shown on the questionnaire with and without the transient Indians whose presence or absence is unrecorded. The hypothesis to be tested in such cases would be whether or not a bias is introduced for the sample results as a function of the transient Indians. A simple means test can be established to test whether or not the mean values of the two distributions are significantly equal from a statistical perspective. 4/

Another assumption in this analysis is that a labor cost of $3.00 per hour will be incurred in the interview and processing phases of the study. Also, a $5.00 per hour charge will be incurred for the administration and analysis stages. The costs of obtaining and analyzing

---

4/ For an example of testing the difference between two means, see the following reference: William Mendenhall, Introduction to Statistics, (Belmont, California, Wadsworth Publishing Company, 1964), p. 151.

Indian statistics can be divided into the preliminary sample stage and the regular sample stage (which is an extension of the preliminary stage to include the total sample selected). The relevant costs can be approximated by the following functions:

$$TC = PC + SC$$
$$PC = F + cN + dn_1 + en_1 + bn_1$$
$$\underline{5}/SC = d\ n_t - n_1 + e\ n_t - n_1 + g(n_t)$$

Where:

- TC = Total costs of planning, obtaining and processing the sample results for a given reservation.
- PC = Costs incurred in the preliminary sample stage.
- SC = Costs incurred to extend the preliminary sample to include a sufficiently large number of units.
- F = Fixed costs associated with such functions as administration, training interviewers, elerical, etc.
- c = Variable cost rate per element associated with dealing with the population elements in the preliminary stage. (i.e., random number assignment and determining a classification strategy if required).
- N = Population size.
- d = Variable cost rate per element of traveling between sample elements.
- n_1 = Preliminary sample size.
- e = Variable cost rate per element of interviewing sample elements and reviewing the results.
- b = Variable cost rate per element of dealing with the preliminary sample elements to estimate the mean income and standard deviation for the determination of an appropriate sample size.
- n_t = Sample size selected from the preliminary study to achieve sufficiently precise results.
- g = Variable cost of processing and analyzing the statistical results recorded on the questionnaire during the interviews.

5/ In certain methods, the preliminary sample may not be fully extended in the regular sampling stage because of the characteristics of a given technique so the cost functions must be altered accordingly.

In order to evaluate each of the four sampling plans, the following organization is used for each method: (1) basic nature; (2) sample size determination; (3) sample program; (4) simulated sample results.

SIMPLE RANDOM SAMPLING

Overview - The most basic form of sampling is simple random sampling, which means that each element (eligible Indian) in the population has an equal chance of being selected for the sample. No attempt is made to separate the population prior to the sample selection phase. Each element in the population is assigned an identification number, a proper sample size is selected, and a table of random digits is used to determine the units to be included in the sample. The interviewer than will visit each of the Indians selected for the sample and ask the various questions recorded on the questionnaire shown in Appendix A. The two major estimators used in this study are the sample mean (an estimate of "how much") and the sample proportion (an estimate of "how many") which are shown in Appendix B, along with their respective standard errors. The sample mean concept is used exclusively to determine an appropriate sample size in this study.

The major advantage of simple random sampling is its simplicity. Also, it is beneficial to consider its basic characteristics, since they provide the foundation for other sampling methods. The population does not have to be divided in any way before the sample is selected so a great

deal of sophisticated analysis of the population characteristics is unnecessary. When the variation among the various Indian incomes and other statistics of interest is narrow, simple random sampling may provide a sufficiently reliable estimator within acceptable cost boundaries and produce results which can be controlled within probabilistic limits. However, an analysis of the standard error with simple random sampling shown in Appendix B reveals its major disadvantage. If a great deal of variance between income levels exists, a sample size which is so large it is prohibitively costly may be required. In such cases, the sampler may be able to achieve more precision per unit of cost with another sampling method. Or geographical dispersion may be so wide that unacceptable costs will be required to achieve satisfactory sample results. In this case, a cluster sampling approach will provide lower costs, although the degree of precision may diminish, so this method must be evaluated carefully to achieve the proper precision-cost tradeoff.

Sample size selection - The standard error should be used to estimate the sample size required for a desired degree of precision. Two approaches are possible to achieve this objective: (1) The standard error or the average individual income can be used to determine an appropriate sample size; (2) The standard error or the proportion for individual income levels can be used for this purpose. As was mentioned

earlier, since so many of the measures of interest are either formally or informally related to the income level of the various Indians, we recommend that either of the standard errors involved with income will provide an effective strategy for the selection of a specific sample size. The resulting sample size can then be used for a given reservation with the assumption that it will provide satisfactory statistical results for all of the measures of interest on the questionnaire. An alternative to this procedure would be to use the same approach for several important estimates listed in the questionnaire and choose a sample large enough that each of these characteristics is estimated with sufficient precision. The first approach, which we employ in this study, assumes that by controlling the standard error of the income measure, a reliable control of the total sampling error is achieved. Because of its effectiveness, we consider only the use of the standard error of the mean individual income estimate for the determination of an appropriate sample size in this study.

Two major decisions must be made in order to determine an appropriate sample size: (1) First, we must determine the amount of error we can tolerate for the income estimate of interest; (2) Second, we must determine the confidence we must have that the estimate will result within the acceptable error limits. The first factor defines the desired precision of the estimate and the second designates the reliability

of achieving the precision objective. We can summarize our dual concern for the precision-reliability combination by looking at the question of, "What is the probability that the mean individual Indian income will be accurate within a given range." This analysis can be expressed as follows:

$$A = kv$$

Where:

A = Precision required for the income estimate.
k = The number of standard deviations expressing the confidence the sampler must have in the estimator. The number 2 often is used to approximate a 95 percent confidence interval. Other possibilities are shown in Appendix D. With k = 2, we know that 95 percent of the time our estimate will be within the acceptable range, A.
v = The standard error defined as the population standard deviation (amount of variance) divided by the square root of the sample size, n.

As was mentioned above, we consider only the use of the standard error of the individual income average for sample size selection purposes. Essentially, the same analysis is valid for the use of the standard error of proportions. A preliminary study should be performed to obtain a reliable estimate of the population standard deviation of individual Indian income. A sample of size 30 would normally be appropriate for this objective. This estimate can then be inserted into the following formula to determine the appropriate sample size which will allow an acceptable error and a satisfactory confidence level:

$$n = \frac{k^2 N S^2}{N A^2 + k^2 S^2}$$

Where:

k = The number of standard deviations necessary to achieve a desired confidence interval for reliability.
N = Population size.
A = Range of error allowed or precision (for example a range of $\pm$ $200) for the average income estimate.
S = The population standard deviation which is normally estimated by the sample variance, s (See Appendix B).

Once the appropriate sample size is determined, the sampler can draw the sample, estimate the measures of interest, and be assured that the error for the mean estimate of individual Indian income will be constrained within acceptable limits.

Sample program - The following steps should be taken to obtain a reliable simple random sample:

1. Identify the population elements consisting of all Indians who are sixteen years or older from the reservation rolls.

2. Assign sequential numbers to the population elements.

3. Obtain a table of random digits from a statistics textbook or the reference listed below. 6/ Determine the number of digits required to represent the population element numbers and their range. For example, in the simulated income information in Appendix C, the numbers 001 to 000 would be used.

4. Decide on a route to use in the table and select a starting point randomly. Select 30 random digits from the table and match them with the Indian population numbers. The 30 individual Indian incomes represent a preliminary sample

---

6/ Acheson J. Duncan, Quality Control and Industrial Statistics, (3rd Edition, Homewood, Illinois, Richard D. Irwin, 1965), p. 938.

which will provide an estimate of the population mean and the population standard deviation which, in turn, will be used to determine an appropriate sample size from the standard error with the formula above.

5. Travel to and interview the 30 Indians selected for the preliminary sample. Use the questionnaire shown in Appendix A to obtain information concerning each of the pertinent questions, including the individual income levels.

6. List the 30 sample incomes. Compute the sample mean and standard deviation for the individual incomes. Use the latter as an estimate for the population standard deviation.

7. Choose an acceptable range of error for the mean, individual income estimate. For example, an error of $\pm$ \$50, $\pm$ \$100, $\pm$ \$200, $\pm$ \$300, etc. might be selected. Also, an acceptable percentage of the time this error will occur must be chosen. Normally, a confidence interval of 95 percent is used for this objective. As the precision is increased by narrowing the acceptable error, the sample size must be increased.

8. Use the formula listed earlier to determine an appropriate sample size of n elements. Extend the preliminary sample of 30 to a sufficiently large number to achieve the required sample size of n elements through the use of a table of random digits.

9. Travel to and interview all Indians chosen for the sample to obtain the information listed on the questionnaire.

10. Tabulate the results of the interviews to quantify the statistical measures of interest for the questionnaire.

11. Compute the new sample mean individual income and the related standard error to evaluate the relevancy of the actual precision and reliability limits incurred.

Simulated Sample Results - In the preliminary stage, a sample of 30 individual incomes was selected from the hypothetical population.

These sample results are shown in Appendix E. The sample mean income was $1,514, with a standard deviation of $2,452. It is obvious that a significant amount of income variation and skewness exists in the simulated population--characteristics which appear likely for actual Indian reservations.

Sample sizes required to achieve specific precision objectives under varying values for the relevant factors are shown in Appendix F. The four factors which must be defined to determine an appropriate sample size are: (1) N, the population size; (2) A, the desired precision; (3) k, the desired confidence level; and (4) v, the standard error. The analysis in Appendix F is based on a confidence level of .95 and the standard deviation estimated from the preliminary sample. Different sample sizes result from varying assumptions concerning N and A. It is obvious from the results of the simulated sampling performance that simple random sampling will be a costly method because of the significant dispersion in the Indian income population. If this situation exists in a real-life case, the costs of simple random sampling required to achieve a satisfactory precision level may be prohibitive. For example, for a reservation of 1,000 eligible Indians, a sample size of 706 elements would be required to insure a sample individual income average within $100 of the true value at the .95 confidence level with the standard deviation of this case. If we want to be within

$50 of the true value, we would have to practically enumerate the entire population, since a sample of 905 elements would be required. Even if we relax the precision requirement and accept the estimator within $200 of the true value 95 percent of the time, a sample of 376 elements would be required. Other combinations are shown in Appendix F for illustrative purposes. While desired precision levels can be attained in this case with simple random sampling, the costs will be high.

The costs shown in Table A are assumed as being realistic for determining the cost of simple random sampling with the hypothetical Indian income population. These costs represent the parameters of the cost functions defined earlier. In actual applications, the specific costs of performing the various functions in the sample survey should be evaluated carefully. It also should be noted that the cost rates involved may change with different sampling plans, as well as economic circumstances.

Table 6

Assumed Costs With Simple Random Sampling

F	=	$100
b	=	$.20
c	=	$.01
d	=	$1.00
e	=	$3.00
g	=	$.80

We assume that a precision level of the sample mean individual income estimate within $200 of the time value is acceptable at the .95 confidence level in simple random sampling. Consequently, a sample size of 376 Indians would be required. Since we have 30 in the preliminary sample, an additional 346 elements are required. The costs involved with the total survey using simple random sampling then will be approximated as follows:

$$
\begin{aligned}
PC &= \$100 + .01\,(1000) + 1\,(30) + 3\,(30) + .20\,(30) \\
&= \underline{\$236} \\
SC &= 1\,(346) + 3\,(346) + .80\,(376) \\
&= \$1,684.80 \\
TC &= \underline{\$1,920.80}
\end{aligned}
$$

Thus, the total cost in this case of interviewing 376 Indians selected with simple random sampling at the desired precision level will be approximately $1,920.80. If a precision of $100 for the income estimate is required, the same cost function would produce approximate costs of $3,504.80. The economic feasibility of these cost levels must be evaluated and compared with that of other sampling plans. After the additional elements have been selected randomly to extend the preliminary sample results to an acceptable sample size, the Indians involved would be interviewed to ask the questions listed on the questionnaire. These statistics then would be processed to describe Indian life based upon a sample which insures a satisfactory precision level concerning

their income performance. Using the final statistical results of Appendix E, we are assured that the true mean individual Indian income will be within the range $1,229 to $1,701 approximately 95 percent of the time. So, with simple random sampling, the potential sampling error is significant.

STRATIFIED SIMPLE RANDOM SAMPLING

When the Indian population has a great deal of variability between the income levels of the various Indians (this appears likely in most cases), a large sample is required to attain a tolerable precision level with simple random sampling and other methods. The previous discussion of simple random sampling illustrates this problem. By reducing the variance, the standard error decreases and higher efficiency (precision per cost unit) can be achieved. The effect of extreme values is offset to a large degree with such action. In order to do so, the objective is to stratify the total population into homogeneous groups. Since individual income is our major statistic of interest for survey design purposes, we would want to group the Indians in such a way that individuals with similar income levels are in the same strata. The total variability of income levels is reduced greatly if this is achieved since the main variance comes from a between group's source because of the homogeneity within the strata. Consequently, an effective precision-cost

trade-off can be achieved since a smaller sample size is required to attain a particular precision objective than with simple random sampling. We still have a probabilistic sample as each element has an equal chance of being selected for the sample.

In order to achieve a significant amount of income homogeneity within the strata, we need to have some reasonable basis with which to classify the population. However, we do not have information about income until the sampling is accomplished and even then we do not know the income levels of all Indians in the population. Thus, we need to obtain information concerning a known measure which is correlated highly with the income performance. This measure can then be used to stratify the population in an appropriate manner. One suggestion for accomplishing this objective is to use vehicle registration statistics. By assuming a high correlation between the age of the vehicle owned by a given Indian and his income level, the registration statistics can be used to classify the population into meaningful strata. If no vehicle is owned, a zero or low income level may be assumed.

In an illustration of stratified simple random sampling with the hypothetical Indian population of Appendix C, we know the various income levels so we can stratify based on our analysis of the income values. In a real-life case, however, we would have to make certain assumptions about the relationships between a correlated variable, such as vehicle registration and income levels. An example of this approach is listed

in Table B. Since perfect correlation is not required to receive significant benefits from stratified sampling, the procedure offers potential advantages.

Table 7

Assumed Correlation Between
Vehicle Registration and Indian Income

Vehicle Age (Years)	Income (Dollars)
None owned or greater than 10	$0-500
6-10	501-1000
4-5	1001-3000
2-3	3001-6000
Less than 2	Greater than 6000

Sample Size Selection - The selection of a sample size that will provide the desired precision results with stratified simple random sampling is based upon the same approach as that illustrated earlier for simple random sampling. Once the various strata are identified, the number of Indians that must be selected to achieve the desired interview results is based on the same basic formula:

$$A = kv$$

Where:

- A = The desired precision of the estimate in terms of the deviation allowed.
- k = The number of standard deviations expressing the confidence the sampler must have in the estimate.
- v = The standard error which is defined in Appendix B for stratified simple random sampling.

Once again, in order to be able to select an appropriate sample size, we need to have a reliable estimate of standard deviation. But since the population is stratified, we need an estimate of the standard deviation from each of the strata. A preliminary sample of 30 Indians from each of the strata. A preliminary sample of 30 Indians from each stratum should be obtained for this purpose. Then, the following formula can be used to select a sample size which offers an acceptable precision level:

$$n_h = \frac{\sum^{L} N_h S_h \quad (N_h S_h)}{N^2 \frac{(1_A)^2}{k} + \quad (N_h S_h^2)}$$

Where:

n_h = Appropriate sample size for the hth stratum.
N_h = Population size for the hth stratum.
S_h = Standard deviation for the hth stratum.
N = Total population size.
k = Number of standard deviations required for an acceptable confidence level.
A = Desired precision describing the range within which the estimate must fall.
L = Number of strata.

Once an appropriate sample size is determined for each stratum, the sampler will draw the sample, estimate all measures of interest from the questionnaire, and be insured that the error for the estimate

of the mean individual Indian income will be within acceptable limits. If the stratification is performed effectively, a smaller sample will be required to achieve comparable precision than that obtained with simple random sampling.

Sample Program

The following steps should be taken to obtain a reliable stratified simple random sample:

1. Identify the population elements consisting of all Indians who are sixteen years or older from the reservation roll.

2. Assign sequential numbers to the population elements.

3. Stratify the population into several classes with the objective of grouping similar income levels together. Use some measure, such as vehicle registration for stratification purposes. Four or five strata normally will be sufficient to achieve the desired results in this type of study. The population element should be matched with vehicle registration statistics according to decision rules, such as those shown in Table B.

4. Randomly select 30 elements from each stratum by using the same random digit procedure discussed earlier for simple random sampling. If less than 30 elements are included in a stratum, use all of them.

5. Travel to and interview each of the Indians selected for the stratified sample. Obtain information related to each question listed on the questionnaire, including the measure of his individual income.

6. Compute the total sample mean individual income and the standard deviation for each stratum. Use the sample standard deviation as an estimate of the population standard deviation for each stratum.

7. Choose an acceptable range of error for the mean individual income estimate (precision). For example, as before an error of ± $50, ± $100, ± $200, etc. might be selected. Also, an acceptable percentage of the time this error will occur must be chosen (reliability).

8. Use the formula given earlier to determine the appropriate sample size for each stratum. Extend the preliminary sample of 30 per stratum, if necessary, to achieve the required sample of n items where n is equal to the sum of the elements drawn from the strata.

9. Interview all Indians chosen for the sample to obtain the information listed on the questionnaire.

10. Tabulate the results of the interviews to quantify the statistical measures of interest from the questionnaire.

11. Compute the new sample mean individual income and related standard error to determine the relevancy of the actual precision and reliability incurred.

Simulated Sample Results - The hypothetical Indian population was stratified as shown in Table C to reduce the effect of the skewed income distribution and related variance concerning the sampling process. A preliminary sample of 30 elements were selected randomly from each of the four strata. The statistical results of the stratified sample process are summarized in Appendix G. Note that the standard error of the average income estimate has been reduced significantly, thus suggesting that a much smaller sample size will be required than with simple random sampling to achieve comparable results. Using the formula for sample size selection with stratified simple random sampling, we confirm this premise. In this case, the preliminary sample itself

will produce a sufficient number of elements for a precision requirement of $100 or more. If we want to narrow the tolerable range to $50, we need only extend the preliminary sample to a total of 191 elements distributed as shown in Table D.

We should keep in mind that even though a sampling error of $50 on the average seems small, this produces a potential error of $50,000 for the total income of a reservation as small as the one with which we are dealing--given the assumed income data. Or, looking at another way, with a mean individual income of about $1,500, $50 represents approximately 3.3 percent of the average income.

Table 8

Strata Established
for 1000 Simulated Individual
Indian Incomes

Stratum	Income Range	# of Elements	
1	$0-999	606	0001-0606
2	1000-2999	235	0607-0842
3	3000-6999	120	0842-0962
4	7000 and above	39	
Total		1000	

Table 9

Sample Size Per Stratum

Stratum	Sample Size D = 50	Sample Size D = 100 or more
1	48	30
2	63	30
3	56	30
4	24	30
Total	191	120

A precision level of $100 would be $100,000 for the income total or 6.6 percent of the average income. If we use the precision limits of $100 as a comparative base, we see that the same results can be achieved with a sample of 120 elements, using stratified simple random sampling that would require 706 units with simple random sampling. Thus, as long as the costs of preparing the population for stratified sampling are not exceedingly high, the stratified method offers substantial cost benefits with acceptable precision limits. We use the assumed costs shown in Table E for an analysis of the costs incurred to obtain a stratified random sample in this case. Once again, since we are dealing presently with a hypothetical case, these costs are based upon expected conditions which should generally prevail. In a specific real-life case, the costs used for analysis purposes should be based on the prevailing conditions.

Table 10

Assumed Costs With Stratified Simple Random Sampling

F = $125
b = $. 30
c = $. 50
d = $1. 00
e = $3. 00
g = $. 80

Some of the individual costs items have changed from those used for simple random sampling because of the basic nature of stratified simple random sampling. We would expect the fixed costs of the preliminary stage to increase since the decision makers will have to spend more time planning an appropriate stratification technique. Also, more processing time will be required for computing the mean and standard deviation, since the stratified sampling formulas are more complicated. Since additional information will be required to achieve the stratification objective (such as vehicle registration statistics), the variable cost of dealing with the population in the preliminary stage will increase substantially. In order to achieve a precision level of $100 for the mean individual Indian income, the following approximate costs would be incurred in this case.

PC = $125 + .50 (1000) + 1 (120) + 3 (120) + .3 (120)
= $1141. 00
SC = $. 80 (120)
96. 00
TC = $1237. 00

Thus, in a situation like this, where a large income variance exists on a given reservation, substantial savings are offered through the use of stratified simple random sampling. Since extreme differences between income levels on a reservation are likely based on previous studies, we strongly recommend stratified simple random sampling as a reliable way to overcome the problem. By incurring a higher cost to prepare the population for sampling through effective stratification, a much smaller sample size will be necessary in cases such as the one illustrated and the precision-cost trade off is much more favorable than that of many other methods. Even in cases where the preliminary sample has to be extended to achieve the desired precision, substantial cost savings should be possible when the wide variation of income is a problem. We can evaluate the validity of the final estimate for the mean individual Indian income as follows: The true value will be within the range of about $1,343 to $1,487 at the .95 confidence level (or saying the same thing, 95 percent of the time). Thus, the precision involved is quite high with stratified simple random sampling.

SINGLE STAGE CLUSTER SAMPLING

Cluster sampling is involved when the Indian reservation is divided into areal units--each of which consist of several Indians (actually, Indian domiciles may be used). In certain cases, single

stage cluster sampling may provide an effective method for applying the Indian questionnaire. The major benefit of cluster sampling is that it is useful where a wide geographical area is involved, since the cost of traveling between domiciles to obtain the sample size required for one of the other methods described in this section may be prohibitive. A secondary benefit of cluster sampling applies to a situation where the population of a specific reservation is not well defined on the population rolls and the interviewer is forced to identify a representative setment of the population so the survey can be taken. By restricting his analysis to localized geographical areas (clusters), the interviewer may be able to avoid attempting to cover an entire population.

Cluster sampling must be used with caution, however, to insure that a satisfactory precision level is attained. This problem can be readily seen from an analysis of the basic characteristics of cluster sampling. Instead of selecting the sample from all segments of the population, the final sample consideration in cluster sampling is directed toward a limited part of the population. We still are dealing with a probability sample, since all elements have an equal chance of being chosen, but the final selection phase is concerned with a limited coverage. In the preliminary stage, the population is divided into groups (clusters) of Indians based upon the geographic conditions involved. These groups (clusters) of Indians located close together

serve as primary sampling units. Recall that in stratified sampling, the objective was to group the Indians in a way that would achieve a great deal of homogeneity within the strata, thus reducing the variability of income levels and necessary sample sizes. In cluster sampling, we usually do not expect similar precision savings, since the geographical conditions primarily dictate the group structures. Within the geographical constraints, the idea is to attempt to attain as much heterogeneity as possible within each of the clusters so they serve as miniature populations. If an extremely large amount of heterogeneity is achieved, it is possible that nominal precision gains compared with other methods may result. But, even though precision gains normally are not achieved, reduced travel costs may offer an attractive precision per cost unit opportunity; since the sample elements are so close together, the sampler can afford to select a larger sample than other methods would allow.

Once the clusters have been identified, the sampling process begins by randomly selecting a sufficient number of clusters that will insure that the desired precision standards are met. Then, each Indian located in the selected clusters is interviewed so the sampler actually enumerates all elements of the chosen clusters. All clusters, therefore, are not chosen for the sample and all of the sampling error is related to this dimension. The cluster method also can be extended to a two stage approach, where we not only sample to choose the clusters,

but also sample within the clusters selected to determine the final group of Indians to be interviewed. We do not cover this case in this study. If each cluster contains Indians with significant differences in their individual income levels, it will serve as a miniature population and the method may produce satisfactory precision results. Since the travel cost of drawing the sample with this method is not related to sample size to the degree other methods are, the precision per unit of cost may be attractive. The two main factors to consider then in evaluating the feasibility of cluster sampling for a given Indian reservation are: (1) the significance of the travel costs; (2) the degree of heterogeneity which can be achieved with the various clusters so the sample size does not have to be extremely large.

Sample Size Selection

The use of the cluster sampling standard error for the determination of a satisfactory sample size is much more complicated and less effective than the processes described for the first two methods. However, since our objective with single stage cluster sampling normally is not to achieve precision gains, a useful approach for determining the selection of an appropriate sample size is to use the standard error for simple random sampling rather than that of cluster sampling. The standard error for simple random rampling is less complicated to work with and provides an effective way to determine a satisfactory sample size with certain limiting assumptions.

Basically, we assume that approximately the same degree of precision incurred with simple random sampling will prevail for cluster sampling. As long as a reasonable amount of heterogeneity is achieved within the clusters, this assumption is realistic. Normally, the sampling error with cluster sampling will be somewhat higher when compared with that of simple random sampling for a given sample size, but if adequate heterogeneity is attained, the difference will not be significant. As was mentioned earlier, modest precision gains may even result if the heterogeneity is sufficiently high. We can use the same approach as that described earlier to determine an appropriate sample size with simple random sampling. A preliminary sample is taken to estimate the population standard deviation and this statistic is combined with the desired precision limit and reliability to calculate an appropriate sample size. The precision results can be evaluated to see if additional sample information should be obtained to further reduce the standard error. Once the total elements have been selected, this number must be translated into clusters, since the latter constitute the sample. Assuming that the same number of elements have been assigned to each cluster, we need only divide the total number of elements by the cluster size to determine the number of clusters which must be selected. For example, if the sample size formula determines that 300 elements must be drawn and each cluster contains 15 elements, 20 clusters would be enumerated for the sample results.

Sample Program - The following steps should be taken to obtain a reliable single stage cluster sample:

1. Identify the population elements consisting of all Indians who are sixteen years or older from the population roll.

2. Match the location of the Indians in No. 1 with a map of the reservation.

3. Based upon an analysis of the geographical conditions shown on the map (for example, villages in existence, spread between domiciles, accessibility, etc.) and an approximation of the income heterogeneity involved, divide the total reservation into several clusters of Indian domiciles. An equal number (or at least approximately so) of eligible Indians should be included in each cluster.

4. Assign sequential numbers to the population elements and separate numbers to the clusters.

5. Obtain a table of random digits and select a simple random sample of 30 Indians, based upon the same procedures discussed in an earlier section.

6. Travel to the location of the Indians selected and interview them with the questionnaire.

7. From the preliminary sample determined in No. 5, calculate the sample mean individual income and related standard deviation. Use the sample standard deviation to determine an appropriate sample size with simple random sampling.

8. Choose an acceptable precision range and reliability limit just as was done with simple random sampling. Use the formula $A = kv$ to calculate an appropriate sample size of n elements.

9. Divide the sample size in No. 8 by the number of elements included in each cluster. The result is the number of m clusters which must be selected.

10. Use the table of random digits to select the number of clusters required for the sample information.

11. Interview all Indians located in the clusters chosen.

12. Tabulate the results of the interviews to quantify the statistical measures of interest from the questionnaire.

13. Compute the new sample mean individual income and the related sample standard error to determine the relevancy of the actual precision and reliability incurred.

Simulated Sample Results

In order to illustrate the application of cluster sampling, we assume that the hypothetical Indian population of 1,000 elements can be divided into 50 clusters of 20 Indians, based upon an assessment of the prevailing geographical conditions and objective of heterogeneity within the clusters. The specific characteristics of a given reservation must be considered carefully in a real-life case to achieve as much heterogeneity of individual Indian income within the clusters as the geographical conditions will allow. But the primary objective is to minimize the travel costs required between the clusters so precision considerations are secondary to the cost factors. A map of the reservation should be used to effectively group the eligible Indians into realistic clusters. The reservation rolls should be matched with the map for this purpose. Local inquiry may be required to supplement, check, and update the information recorded on the map and population roll.

Using the rule suggested earlier for sample size, we find that 19 clusters must be chosen to receive the 376 elements required for an income estimate within $200 of the true value 95 percent of the time. The 19 clusters give us 380 sample elements to work with. The preliminary results would be obtained for the cluster sampling method with simple random sampling to identify the significance of the income variance as was discussed earlier. Then, the proper number of clusters would be selected to obtain the required sample size. The sample results with cluster sampling are shown in Appendix H. The sample mean income was $1,464, with a standard error of $29.43. The approximate costs of obtaining the cluster sample information are based upon the data shown in Table 11.

Table 11

Assumed Costs With Single Stage Cluster Sampling

F	=	$130
d	=	$.20
e	=	$.30
d	=	$.30 for SC; $1.00 for PC
e	=	$3.00
g	=	$.80

Once again, many of the individual cost items have changed when compared with the first two methods because of the unique characteristics of cluster sampling. The fixed costs will increase slightly from those of stratified sampling, since we must be concerned with

geographical factors, as well as a classification strategy which achieves some heterogeneity. The variable cost of processing the preliminary sample in order to determine an appropriate sample size will be equal to that incurred for simple random sampling. The travel cost of obtaining the regular sample will be lower than that of other methods, since we are dealing with clusters. An additional cost is expected with cluster sampling because of the fact that the preliminary sample cannot be extended to the regular sample, since the latter consists of clusters. Some of the preliminary sample units no doubt will be located in the clusters chosen, but we will not know how many there will be until the clusters are chosen.

$$
\begin{aligned}
PC &= 130 + .30\,(1{,}000) + 1\,(30) + .20\,(30) \\
&= 556 \\
7/\,SC &= .30\,(380) + 3\,(380) + .80\,(380) \\
&= 1{,}558 \\
7/\,TC &= \underline{\underline{\$2{,}114}}
\end{aligned}
$$

Therefore, the precision per unit of cost with cluster sampling would be measured by a small standard error of approximately $29.43 at a cost of about $2,114. Consequently, we have a better precision per unit of cost here than we had with simple random sampling when the standard error was $118 and the cost was about $1,921. So, for a

---

7/ It is likely that some of the preliminary sample elements also will be included in the regular sample, so the cost function may be overstated.

slightly higher cost, cluster sampling produces a lower standard error in this case. This outcome resulted from the high degree of heterogeneity with the clusters selected for the sample and will not always occur, since in the typical case, we use cluster sampling for cost benefits but do not expect precision gains. We can analyze the actual precision that results by noting that we can expect the true value of the mean individual Indian income to be within the range of about $1,404 to $1,524, approximately 95 percent of the time. So, again we see that the error range is narrow in this case.

SYSTEMATIC SAMPLING

The final method discussed in this section for conducting the Indian interviews is systematic sampling. The chief advantage of this method is its simplicity to accomplish. A decision rule is built in to inform the sampler as to the elements he must select rather than requiring him to depend on a table of random digits. Also, in certain conditions, the standard error with systematic sampling is less than that incurred with alternative methods--although there is no guarantee of this outcome. The major disadvantage of systematic sampling is the danger of periodicity, which means that elements with essentially the same characteristics are selected regularly so a bias is introduced.

This potential problem is obvious when we analyze the major characteristics of systematic sampling.

The basic idea is that every kth Indian located on the reservation roll is selected for the sample. Once the population of eligible Indians is identified, an appropriate sample size n is selected and k is determined as $\frac{N}{n}$. A popular approach is to begin the sampling by randomly selecting an Indian in the population listed between 1 and k. This provides a starting number and every kth Indian then is selected to constitute an effective sample. For example, in the hypothetical population of 1,000 Indian incomes, we might desire a sample of 200, so k is 5. We would randomly select a number between 1 and 5 to begin the sampling. Assume we select a 3. The sample would include Indians listed in the 3rd, 8th, 13th, 18th, etc. positions. The basic problem relates to the fact that the data involved may be periodic in relation to the order of the listing. The selected elements may be alike and thus would not be representative of the total population. If essentially the same value is repeated over the sample selected, a significant standard error is introduced. An effective way to minimize the possibility of periodicity is to randomize the population roll used to draw the sample. The alphabetized list maintained on the reservation roll would be changed to a randomized order through the use of a table of random digits. Then the systematic sample would be drawn after k was selected.

A counter-balancing disadvantage of randomizing the population is that precision gains are incurred with systematic sampling when there is a high correlation between the income level of adjacent Indians located on the population list. The serial correlations involved should decrease as the interval between the individual Indians increases. When the serial correlations of near-by Indians are large with decreasing correlations between the units farther apart, significant precision gains may result with systematic sampling. Therefore, the sampler can organize the population Indian income data in a way that achieves the serial correlation objective and, at the same time, avoids the threat of periodicity, the randomization procedure should not be used. However, it is doubtful that the information required for such an effective organization will be available, since generating it is the primary objective of the sample process. Thus, the use of systematic sampling in a situation like an Indian study is unlikely to produce consistent precision gains and its results will be similar to those of simple random sampling. Since precision gains probably cannot be achieved intentionally, the major advantage of the method is its simplicity and structured rule for the selection of the sample elements.

Sample Size - The standard error with systematic sampling is not amenable for the selection of an appropriate sample size. However,

since our objective with this method in a case, such as an Indian study, normally is not to achieve precision gains anyway, the standard error with simple random sampling can be used in a preliminary study to determine a satisfactory sample size. The approach described earlier for simple random sampling can be used to define an effective sample size by taking a preliminary survey of 30 Indians randomly, determining the sample standard deviation, and using the formula defined earlier for sample size selection. Once n is established, k also is defined by dividing N by n (Francional results are rounded to the nearest whole number). A random start between 1 and k initiates the sampling process and every kth Indian is selected for the sample. Consequently, we will achieve approximately the same precision-cost results as we would have with simple random sampling, but we have a systematic strategy to follow.

Sample Program - The following steps should be taken to obtain reliable sample information with systematic sampling:

1. Identify the population elements consisting of all Indians who are sixteen years or older from the reservation roll.

2. Randomize the sequential order of the list in No. 1 to minimize the chance of periodicity.

3. Randomly select a preliminary sample of 30 Indians from the new list.

4. Travel to and interview the 30 Indians selected for the preliminary sample. Use the questionnaire shown in Appendix A to obtain information concerning each of the pertinent questions, including the individual income levels.

5. List the 30 sample incomes. Compute the sample mean and standard deviation for the individual incomes. Use the latter as an estimate for the population standard deviation.

6. Choose an acceptable range of error for the mean individual income estimate. Also, determine a satisfactory reliability level--the percentage of the time this error must occur.

7. Use the formula listed earlier (including the precision, reliability, and estimate of the population standard deviation) to determine an appropriate sample size of n elements.

8. Determine k, the sampling interval, by dividing the sample size into the population $(\frac{N}{n})$.

9. Randomly select a starting point between 1 and k. Then choose every kth Indian to constitute a total sample of n items. (Round any uneven number for k to the nearest digit).

10. Travel to and interview all Indians selected for the sample of size n. (Hopefully, some of those selected already will have been interviewed in the preliminary stage--although there is no guarantee of this). Obtain the pertinent information listed on the questionnaire.

11. Tabulate the results of the interviews to quantify the statistical measures of interest from the questionnaire.

12. Compute the new sample mean individual income and related standard error to evaluate the relevancy of the actual precision and reliability limits incurred.

Simulated Sample Results - The same preliminary sample results generated with simple random sampling earlier (Appendix E) were used to determine an appropriate systematic sample. Precision and reliability

limits of $200 and 95 percent, respectively, were assigned for the systematic sampling project. A sample of 376 elements were required to achieve these objectives, so k was set equal to 3 to provide approximately the required number of elements. A random selection of the number 2 established the starting point and every third number in the randomized version of the hypothetical Indian population was selected (2nd, 5th, 8th, 11th, 14th, etc.). The sample of 333 elements produced a mean individual Indian income of $1,448 and a standard error of approximately $125, as shown in Appendix I. The major advantage of this method obviously is its simplicity, since an automatic decision rule is built in to determine the elements required for the sample.

The cost estimates with systematic sampling are shown in Table 12. These costs are very similar to those incurred with simple random sampling, except for the variable costs of preparing for the preliminary sample.

Table 12

Assumed Costs With Systematic Sampling

F	=	$100
b	=	$.20
c	=	$.05
d	=	$1.00
e	=	$3.00
g	=	$.80

Based upon these costs, the approximate costs with systematic sampling would be as follows:

$$
\begin{aligned}
PC &= \$100 + .05\,(1000) + 1\,(30) + 3\,(30) + .2\,(30) \\
&= 276 \\
8/\ SC &= 1\,(333) + 3\,(333) + .80\,(333) \\
&= 1,298.40 \\
8/\ TC &= \underline{1,588.40}
\end{aligned}
$$

The costs with systematic sampling and the hypothetical Indian population are lower than they are with simple random sampling, but the standard error is higher. However, a smaller sample has caused this lower cost because of the rounding process to compute k. Since the sample is smaller, we have not achieved the precision limits originally planned, but the effect of the difference will be small. We can rely on the fact that the true value for the mean individual income will be in the range of $1,178 to $1,718 at the .95 confidence level. Thus, we have controlled the standard error within reasonable limits, although they are wider than with stratified and cluster sampling.

SUMMARY

Four basic sampling methods which will provide objective and reliable ways to select an adequate number of Indians from the population have been discussed in this section. The use of one of these sample designs eliminates the need to deal with the entire Indian population

involved but does so with known precision and reliability limits. A hypothetical Indian population was employed to illustrate the use and hypothetical cost of the various plans.

Sampling with replacement is assumed throughout the analysis because of the nature and objectives of the Indian interviews. The focus to control the precision and reliability of the sample results is on the individual Indian incomes, since so many of the statistical measures of interest on the questionnaire are highly correlated with the individual income levels. By controlling the sampling error for the income estimate, the authors suggest the total errors of sampling can be regulated effectively.

Each of the sampling method discussed have certain advantages and disadvantages--primarily from a cost and precision perspective. As was mentioned earlier, we suggest strongly that stratified simple random sampling will offer significant gains for most Indian studies. The diverse income levels of a reservation are likely to create a significant amount of variation and skewness, which will demand a large sample size with simple random sampling, cluster sampling, and systematic sampling. Precision gains will be achieved if an effective system is designed to structure the total population into strata. A known statistic which is highly correlated with the individual Indian income levels can be used to achieve an effective stratification strategy. Vehicle registration statistics, while they may not be correlated perfectly with incomes,

---

8/ Once again, it is likely that some of the preliminary sample elements also will be included in the regular sample so the cost function may be overstated.

should provide a sound basis for the stratification process. Since the total variance is reduced primarily with stratification to a between group's variance because of the homogeneity within each stratum, a smaller sample is necessary to achieve given precision objectives than that required with other methods. Consequently, the costs are reduced substantially. Once a reasonable sample size is determined, the appropriate number of Indians are interviewed through the use of the questionnaire shown in Appendix A to investigate the relevant facts concerning Indian life. As long as one of the four methods discussed in this section is selected and employed, we are assured that the statistical results will be based upon an objective foundation, subject to a controlled sampling error.

APPENDIX A

Sample Questionnaire for Indian Demographic Studies

Name _____
Address _____
_____
Sex _____
Year of birth _____
Code _____
Date Interviewed _____
Interviewer _____

ASK ALL RESPONDENTS

1. Did you work last week? Yes_____ No _____

2. Did you work during the past month? Yes _____ No _____

3. How many hours a week do you usually work at your job?

1. 1-14 _____
2. 15-29 _____
3. 30-34 _____
4. 35-39 _____
5. 40 _____
6. 41-48 _____
7. 49-59 _____
8. 60 or more _____

4. If you usually work less than 35 hours per week, what is the reason?

1. Slack work _____
2. Material shortage _____
3. Plant or machine repair _____
4. Could only find part time work _____
5. Labor dispute _____
6. Bad weather _____
7. Illness _____
8. Too busy with housework, school, business, personal _____
9. Did not want full time work _____
10. Full time work week under 35 hours _____
11. Other reasons _____

5a. Do you usually work

1. Year round _____
2. Seasonal _____
3. Irregular _____

5b. If you work seasonal, do you usually work in

1. Spring _____
2. Winter _____
3. Fall _____
4. Summer _____

6. What were you doing most of last month?

1. Working _____
2. At job but not working _____
3. Looking for work _____
4. Keeping house _____
5. Going to school _____
6. Not able to work _____
7. Retired _____
8. Other _____

7. Is your present job located on the reservation? Yes _____ No _____

8. Where did you learn your present job?

1. Taught by employer _____
2. Government training program _____
3. Armed Services _____
4. Formal schooling _____
5. Other _____

Questions for Those Not Usually Working

9. Have you looked for a job during the past month? No _____ Yes _____

If yes, full, part time, or both?

10. If looking for work, what have you been doing in order to find work?

1. BIA _____
2. State Employment Service _____
3. Private employment service _____
4. Checked with employer _____
5. Friends or relatives _____
6. Placed or answered ads _____
7. Other _____

5. Health problems _____
6. Personal problems _____
7. Lack of transportation _____
8. Do not want to leave reservation _____
9. Other _____

12. If you are not looking for work, why?

1. Believes no work is available _____
2. Could not find work _____
3. Lack of schooling, training, or experience _____
4 Employer thinks too young or too old _____
5. Cannot arrange for child care _____
6. Do not want to leave reservation _____
7. Family responsibilities _____
8. In school _____
9. Ill health _____
10. Other _____

13. When did you last work at a regular full or part time job?

1. Within past 12 months _____
2. Within last 2 years _____
3. Between 2 and 3 years _____
4. 3 - 4 years _____
5. 4 - 5 years _____
6. 5 or more years _____
7. Never worked _____

14. Why did you leave your last job?

1. Personal reasons _____
2. Ill health _____
3. Retirement _____
4. Seasonal job completed _____
5. Laid off _____
6. Unsatisfactory work conditions _____
7. Other _____

11. Why do you think you have had difficulty getting a job?

1. No jobs available in line your work _____
2. Age - too young or too old _____
3. Lack of necessary skills or experience _____
4. Lack of necessary education or training _____
5. Health problems _____
6. Personal problems _____
7. Lack of transportation _____
8. Do not want to leave reservation _____
9. Other _____

12. If you are not looking for work, why?

1. Believes no work is available _____
2. Could not find work _____
3. Lack of schooling, training, or experience _____
4. Employer thinks too young or too old _____
5. Cannot arrange for child care _____
6. Do not want to leave reservation _____
7. Family responsibilities _____
8. In school _____
9. Ill health _____
10. Other _____

13. When did you last work at a regular full or part time job?

1. Within past 12 months _____
2. Within last 2 years _____
3. Between 2 and 3 years _____
4. 3-4 years _____
5. 4-5 years _____
6. 5 or more years _____
7. Never worked _____

14. Why did you leave your last job?

1. Personal reasons _____
2. Ill health _____
3. Retirement _____
4. Seasonal job completed _____
5. Laid off _____
6. Unsatisfactory work conditions _____
7. Other _____

15. Describe your present job or your last job (within 5 years)

_____ Has not worked in 5 years (skip rest of section)

16. What is the name of your present or last employer?

17. What kind of business or industry is this?

17a. What kind of work were you doing?

18. Does this job involve working
(check one or more)

1. In an office _____
2. Inside, but not in an office _____
3. Outside _____
4. Machinery (specify) _____
5. As a helper _____

SIC code 51-53 _____

DOT code 54-59 _____

19. What other types of employment have you had in the past five years?

Employer	Occupation	Code	Where Learned*	
_____	_____	_____	_____	60-63___
_____	_____	_____	_____	64-67___
_____	_____	_____	_____	68-71___

*Place of Learning

1. Employer _____
2. Government training program _____
3. Armed Services _____
4. Formal education _____
5. Other _____

20. Have you been trained for any type of work not listed above? 72___
1 no 0 yes (specify type of training and approximate date)

21. Do you hold a union card? 73___
1 yes 0 no

22. Where does your family usually purchase groceries? 74___

1 On the reservation
2 Off the reservation
3 Half and half
4 Don't know

22a. How does your family usually pay for groceries at the time of purchase: 75___

1 Cash 2 Credit 3 Don't know

23. Where does your family usually purchase automobile repairs? 76___

1 On the reservation
2 Off the reservation
3 Half and half
4 Don't know
5 None

Zone 77-78___
Subsample 79___

Card 80___
8___

23a. How does your family usually pay for automobile repairs?

1 Cash
2 Oil company credit card
3 Other credit
4 Don't know

24. Where does your family usually purchase clothing: 9___

1 On the reservation
2 Off the reservation
3 Half and half
4 Don't know

24a. How does your family usually pay for clothing? 10___

1 Cash 2 Credit 3 Don't know

25. Are you now enrolled in a school or training program: 11___

1 yes specify _____
0 no

26. What is the highest grade of school you have completed? 12-13___

01	04	07	10	13	16
02	05	08	11	14	17 or more
03	06	09	12	15	

27. Do you have a high school diploma or a GED 14___

1 yes 0 no

28. Do you have any

College degrees
1 yes (specify) 15___
0 no

Technical or vocational school certificates
1 yes (specify) 16___
0 no

29. What language is spoken most frequently in your home? 17___

1 An Indian language 2 English 3 Other

30. Do you speak English? 18___

1 yes 0 no

31. Do you read English? 19___

1 yes 0 no

32. What is your marital status? 20___

1 Married 2 Divorced 5 Never married
3 Widowed 4 Separated

33. How many children have you? 21___

1 One 4 Four 7 Seven
2 Two 5 Five 8 More than 7
3 Three 6 Six 9 None

34. How old are your children?

____ ____ ____ ____ ____ ____ ____ ____ 22___

35. What were the sources of income received by you in the last twelve months?

1 yes 0 no Gifts from children or relatives or churches . 23___
1 yes 0 no From sale of handicrafts..... 24___
1 yes 0 no Self-employed income (business, farm trade, or professional enterprise) individual or as a partner..... 25___
1 yes 0 no Earnings from the ownership of a farm or ranch, a craft shop or retail store, a fishing, tourist, or guide-service business.. 26___
1 yes 0 no Earnings from a trade (carpenter, well-driller, paying contractor, auto mechanic, barber, bookeeper, ranch manager, dress-maker, nurse, etc.)..... 27___
1 yes 0 no Pension (s) (specify)_____.... 28___
1 yes 0 no Assistance payments from Bureau of Indian Affairs..... 29___

1 yes	0 no	Assistance payments from other public or private sources	30___
1 yes	0 no	Interest or dividends on personal loans and investments	31___
1 yes	0 no	Money receipts from tribal or individual allottee sources (timber sales, leases, permits, royalties, annuity payments, dividend payments)	32___
1 yes	0 no	Judgment or settlement funds	33___
1 yes	0 no	Sale of property	34___
1 yes	0 no	Veteran's payments	35___
1 yes	0 no	Social Security benefits	36___
1 yes	0 no	Unemployment insurance	37___
1 yes	0 no	Other (specify) _____	38___
1 yes	0 no	None	39___

36. What was <u>your</u> total money income in 1967? 40___

1	$	0		5	$ 2,000 - 2,999
2	$	1	499	6	$ 3,000 - 4,999
3	$	500 -	999	7	$ 5,000 - 9,999
4	$	1,000 -	1,999	8	$10,000 and over

37. Did <u>you</u> receive any non-money income such as

1 yes	0 no	Homegrown and consumed agricultural products	41___
1 yes	0 no	Homemade clothing	42___
1 yes	0 no	Goods exchange for other goods	43___
1 yes	0 no	Other	44___

38. What is the monetary equivalent of <u>your</u> additional income? 45___

1 $ 0
2 $ 1 - 499
3 $ 500 - 999
4 $1,000 or over

39. What were the sources of income received by your family in the last twelve months:

1 yes 0 no	Gifts from children or relatives or churches	46___
1 yes 0 no	From sale of handicrafts	47___
1 yes 0 no	Self-employed income (business, farm trade, or professional enterprise) individual or as a partner	48___
1 yes 0 no	Earnings from the ownership of a farm or ranch, a craft shop or retail store, a fishing, tourist, or guide-service business	49___
1 yes 0 no	Earnings from a trade (carpenter, well-driller, paying contractor, auto mechanic, barber, ranch manager, dressmaker, nurse, etc.)	50___
1 yes 0 no	Pension(s) (specify)_____	51___
1 yes 0 no	Assistance payments from Bureau of Indian Affairs	52___
1 yes 0 no	Assistance payments from other public or private sources	53___
1 yes 0 no	Interest or dividends on personal loans and investments	54___
1 yes 0 no	Money receipts from tribal or individual allottee sources (timber sales, leases, permits, royalties, annuity payments, dividend payments)	55___
1 yes 0 no	Judgment of settlement funds	56___
1 yes 0 no	Sale of property	57___
1 yes 0 no	Veteran's payments	58___
1 yes 0 no	Social Security benefits	59___
1 yes 0 no	Unemployment insurance	60___
1 yes 0 no	Other (specify)_____	61___
1 yes 0 no	None	62___

40. What was the total monetary and non-monetary income of your family in 1967? 63___

1	$ 0 - 499	5	$ 3,000 - 4,999
2	$ 500 - 999	6	$ 5,000 - 9,999
3	$1,000 - 1,999	7	$10,000 and over
4	$2,000 - 2,999		

Sex 1M 0F	64___
Year of birth	65-66___
Zone	77-78___
Subsample	79___
Card	80 2___

Health Planning

1. How do you describe your general state of health?

   excellent_____
   good_____
   fair_____
   poor_____

2. Do you have a disability which prevents you from working?

   yes_____ no_____

3. Have you ever visited a doctor?

   yes_____ no_____

4. Have you ever used the Indian Health Service Clinic?

   yes_____ no_____

5. Do you prefer to use?

   Indian Health Service_____
   Off reservation private facilities_____

6. If you visit the Indian Health Service, how often during a year?

   1_____ 3_____ 5_____ 7_____ 9_____
   2_____ 4_____ 6_____ 8_____ 10_____

7. If female and have had children, were they born

   in a hospital on the reservation_____
   clinic facility on the reservation_____
   at home_____

8. Were you attended?

   by a physician_____
   mid-wife_____

9. How many visits did you make to a physician prior to birth?

1_____ 5_____ 9_____
2_____ 6_____ more than 9_____
3_____ 7_____
4_____ 8_____

10. Did you take your children to a physician for a check up after birth?

yes_____ no_____

11. If yes, was this a routine check up or was it because of illness?

Explain_____

12. Have you had the following immunization and the approximate dates:

smallpox_____
diptheria_____
tetanus_____
measles_____
typhoid_____

13. What is the source of your water supply?

well_____
local water company_____
other_____

Housing Needs

Questions to be asked of heads of households

1. Do you own your own home?

   yes_____ no_____

2. Do you rent?

   yes _____ no _____

3. How old is the home in which you live?

1-3 years _____	11-15 years _____	26-30 years _____
4-6 years _____	16-20 years _____	over 30 years _____
7-10 years _____	21-25 years _____	

4. How many rooms in your house?

1_____	4_____	7_____
2_____	5_____	8_____
3_____	6_____	9_____

5. Is your bathroom?

   1 inside_____
   2 outside_____

6. Do you have?

   1 hot water
   2 cold water

7. How many people live in your home permanently?

1_____	4_____	7_____
2_____	5_____	8_____
3_____	6_____	9_____

8. Do others live in your home temporarily?

yes_____ no_____

9. What is the relation to you of those who live with you temporarily?

1 Father_____
2 Mother_____
3 Older Son & Family_____
4 Older Daughter & Family_____
5 Other relatives (specify)_____
6 Friends_____

10. How far is your house from where you work?

Less than one mile_____
1-2 miles_____
3-5 miles_____
6-8 miles_____
9-10 miles_____
More than 10 miles_____

11. What type of heat do you use in your house?

gas stove _____
electric stove_____
coal stove_____
wood_____

Labor Force Information

ASK ALL RESPONDENTS

1. What were you doing most of last year? 8___

   1. Working
   2. With a job but not a work
   3. Looking for work
   4. Keeping house
   5. Going to school
   6. Unable to work
   7. Retired
   8. Other (specify)_____

2. How much did you work last year? 9___

   1. 0 (go to green)
   2. 1-3 months (go to pink)
   3. 4-6 months (go to pink)
   4. 7-9 months (go to pink)
   5. 10-12 months (go to pink)

APPENDIX B

Notation and Formulas
Used for Sample Survey Designs

Simple Random Sampling

$\bar{x}$ = Estimate of mean Indian characteristic, such as individual income

$$\bar{x} = \frac{\sum_{n} xi}{n}$$ [1/]

Where: xi = ith element selected for the sample
n = sample size

v = The standard error of the sample estimate of a mean characteristic, such as income

$$v^2 = \frac{(N-n)}{N} \frac{S^2}{n}$$

Where: N = population size
S^2 = variance of the population
$S^2 = \frac{\sum N(Xi-\bar{X})^2}{N}$ & is estimated by s^2, the sample variance
($\bar{X}$ is the true population mean)

$$s^2 = \frac{\sum_{n} (xi-\bar{x})^2}{n-1}$$

xi = ith sample element
$\bar{x}$ = sample mean

p = Estimate of proportion of Indians exhibiting a particular characteristic, such as a given income level.

p = x/n, where x = total Indians in the sample with the characteristic

v_p = The standard error of the estimate of proportion

$$vp^2 = \frac{(N-n)}{N} \frac{pq}{n}$$

Where: p = the estimate of the proportion exhibiting the characteristic of interest
q = the estimate of the proportion not exhibiting the characteristic of interest

---

1/ $\sum$ means to sum over all relevant values.

x = Estimate of the population total for a particular characteristic, such as total individual income.

$x = N\bar{x}$ or Np

v_t = The standard error of the sample total.

$$v_t^2 = N^2 \frac{(N-n)}{N} \frac{S^2}{n}$$

Stratified Simple Random Sampling

$\bar{x}$ = Estimate of mean Indian characteristic, such as individual income.

$$\bar{x} = \frac{\sum_{L} N_h \bar{x}_h}{N}$$ Where: N_h = Population elements in hth stratum

$\bar{x}_h$ = Sample mean in hth stratum

$$\bar{x}_h = \frac{\sum^{n_h} x_h i}{n_h}$$ $x_h i$ = ith element in hth stratum

L = Number of strata

v = The standard error of the sample estimate of a mean characteristic, such as income.

$$v^2 = \frac{1}{N^2} \sum^{L} N_h^2 \frac{(N_h - n_h)}{N_h} \frac{S_h^2}{n_h}$$ Where: S_h^2 = Variance of the elements in the hth stratum.

$$S_h^2 = \frac{\sum^{N_h} (X_h i - \bar{X}_h)^2}{N_h}$$ & is estimated by s_h^2, the sample variance.

$$s_h^2 = \frac{\sum^{n_h} (x_h i - \bar{x}_h)^2}{n_h - 1}$$ $x_h i$ = sample elements in hth stratum

$\bar{x_h}$ = sample mean of hth stratum

Cluster Sampling (Single Stage)

$\bar{x}$ = Estimate of mean Indian characteristic, such as individual income.

$$\bar{x} = \frac{\sum^{n} x_i}{n}$$

Where: x_i = ith element selected from enumeration of the clusters chosen for the sample.

n = number of elements included in the sample from m clusters sampled from total available.

v = The standard error of the sample estimate of mean characteristic, such as income.

$$v^2 = \frac{M^2}{N^2} \frac{(M-m)}{\frac{M}{m}} Sc^2$$

Where: M = number of clusters in population

m = number of clusters selected for sample.

$$Sc^2 = \frac{\overset{M}{(Xc-X/M)^2}}{M-1}$$

X_c = total value of elements of cth cluster.

X = total value of elements of all clusters.

S_c^2 is estimated with s_c^2.

$$s_c^2 = \frac{\sum^{m} (x_c - x/m)^2}{m-1}$$

x_c = total value of elements of cth cluster sampled.

x = total value of elements of all clusters sampled.

Systematic Sampling

$\bar{x}$ = Estimate of mean Indian characteristic, such as individual income.

$$\bar{x} = \frac{\sum^{n} xi}{n}$$

Where: xi = An individual element selected for the sample.

v = The standard error of the sample estimate of a mean characteristic, such as income.

v^2 = Approximated as $\frac{(N-n)}{N} \frac{S^2}{n}$

S^2 is estimated by s^2

$$s^2 = \sum \frac{(xi-\bar{x})}{n-1}^2$$

APPENDIX C
1000 Hypothetical Individual Indian Incomes Simulated From Indian Manpower Resources In The Southwest Study

Row/Column

	1	2	3	4	5	6	7
1	0	0	0	0	57	42	180
2	0	0	0	0	33	57	141
3	0	0	0	0	5	73	169
4	0	0	0	0	92	13	168
5	0	0	0	0	54	62	189
6	0	0	0	0	41	99	174
7	0	0	0	0	83	60	137
8	0	0	0	0	84	49	182
9	0	0	0	0	12	11	157
10	0	0	0	0	81	187	193
11	0	0	0	0	59	104	102
12	0	0	0	0	18	107	104
13	0	0	0	0	79	177	124
14	0	0	0	0	45	188	165
15	0	0	0	0	13	113	172
16	0	0	0	0	91	151	131
17	0	0	0	0	24	148	174
18	0	0	0	0	94	124	197
19	0	0	0	0	22	157	164
20	0	0	0	0	82	127	127
21	0	0	0	0	62	189	121
22	0	0	0	0	89	181	165
23	0	0	0	0	81	115	161
24	0	0	0	0	43	199	181
25	0	0	0	0	60	180	180
26	0	0	0	0	93	191	166
27	0	0	0	0	42	152	109
28	0	0	0	60	12	179	151
29	0	0	0	34	71	166	193
30	0	0	0	4	9	130	137
31	0	0	0	9	79	115	119
32	0	0	0	42	41	186	191
33	0	0	0	49	73	154	137
34	0	0	0	9	92	167	246
35	0	0	0	69	63	183	202
36	0	0	0	82	97	102	226
37	0	0	0	30	64	142	282
38	0	0	0	86	86	195	241
39	0	0	0	83	62	186	290
40	0	0	0	43	42	145	242

Appendix C con't.

	8	9	10	11	12	13	14
1	228	272	327	470	434	593	733
2	204	219	367	484	463	544	745
3	206	267	379	450	416	520	743
4	209	259	317	483	499	513	713
5	239	287	377	435	474	538	713
6	239	240	335	402	428	562	702
7	250	267	318	481	410	582	793
8	203	221	302	443	405	585	781
9	231	244	370	421	428	530	766
10	284	206	357	408	401	631	725
11	288	218	327	482	411	642	721
12	257	205	320	455	477	669	756
13	259	222	348	413	488	611	710
14	218	296	323	458	406	642	769
15	284	287	354	462	496	616	784
16	278	318	328	458	422	699	771
17	234	381	346	426	465	677	770
18	296	386	305	442	401	615	783
19	201	379	391	481	465	670	774
20	272	303	367	479	542	630	746
21	247	327	344	433	511	625	747
22	268	335	326	439	566	659	763
23	233	348	364	445	549	690	738
24	288	381	378	426	540	656	786
25	207	340	362	497	579	622	719
26	220	338	317	424	531	694	795
27	232	320	304	496	524	665	729
28	238	362	381	409	582	676	795
29	284	379	303	438	510	694	775
30	252	390	326	480	537	634	846
31	204	368	358	414	560	643	856
32	203	333	369	464	557	620	807
33	262	372	352	422	544	679	893
34	298	367	347	482	552	660	888
35	258	368	305	405	506	657	841
36	287	379	382	423	594	657	825
37	266	314	343	495	543	605	894
38	279	321	484	419	528	671	842
39	247	375	476	469	536	639	879
40	273	392	465	425	562	769	862

Appendic C con't.

	15	16	17	18	19	20	21
1	804	988	1157	1401	1681	1899	2309
2	808	916	1123	1490	1628	1887	2304
3	889	953	1265	1425	1635	1862	2392
4	852	970	1262	1442	1625	1826	2312
5	871	973	1272	1427	1653	1862	2308
6	860	970	1266	1467	1618	1859	2377
7	878	1085	1289	1443	1643	1878	2426
8	814	1062	1257	1404	1687	1888	2450
9	875	1030	1226	1438	1622	1946	2443
10	891	1070	1225	1477	1688	1923	2415
11	865	1030	1272	1478	1670	1928	2403
12	858	1005	1231	1429	1652	1983	2480
13	862	1097	1268	1483	1768	1979	2543
14	816	1032	1274	1493	1747	1905	2583
15	836	1037	1267	1417	1778	1975	2524
16	888	1037	1244	1446	1760	1994	2588
17	849	1007	1289	1515	1706	1940	2515
18	861	1091	1270	1591	1702	1928	2591
19	835	1041	1241	1562	1715	1925	2662
20	934	1092	1289	1584	1715	1954	2684
21	994	1009	1384	1545	1761	1948	2645
22	976	1022	1345	1555	1798	1932	2655
23	965	1076	1328	1563	1763	2014	2663
24	994	1048	1332	1590	1780	2044	2690
25	922	1140	1387	1523	1746	2006	2723
26	990	1121	1389	1556	1753	2001	2756
27	933	1159	1355	1502	1735	2074	2702
28	960	1178	1394	1547	1745	2087	2747
29	979	1185	1398	1529	1732	2113	2729
30	920	1144	1378	1550	1777	2125	2750
31	943	1180	1333	1532	1852	2119	2832
32	943	1173	1398	1593	1812	2165	2893
33	971	1135	1377	1517	1813	2177	2817
34	939	1187	1364	1568	1888	2174	2868
35	905	1185	1398	1641	1895	2207	2888
36	971	1194	1346	1607	1820	2213	2817
37	957	1178	1362	1678	1887	2231	2965
38	972	1179	1333	1641	1893	2242	2927
39	930	1149	1403	1651	1852	2276	2935
40	982	1109	1468	1606	1864	2263	2969

Appendix C con't.

	22	23	24	25
1	2965	3776	4553	6991
2	3096	3824	4607	7216
3	3095	3851	4622	7137
4	3086	3895	4689	7004
5	3083	3800	4636	7293
6	3022	3877	4630	7389
7	3165	3961	4780	7429
8	3147	3911	4701	7525
9	3180	3946	4781	7592
10	3163	3929	4744	7524
11	3112	3931	4731	7673
12	3284	4031	4804	7745
13	3257	4052	4817	7833
14	3222	4012	4869	7928
15	3257	4013	4807	8086
16	3212	4088	4857	8141
17	3399	4152	4940	8197
18	3373	4193	4995	8240
19	3388	4187	4918	8372
20	3354	4129	5023	8497
21	3301	4195	5167	8591
22	3483	4299	5184	8688
23	3486	4287	5243	8760
24	3406	4262	5329	8838
25	3404	4226	5453	8979
26	3442	4262	5502	9051
27	3583	4359	5605	9143
28	3525	4378	5745	9297
29	3540	4388	5847	9441
30	3576	4321	5960	9322
31	3522	4323	6082	9583
32	3665	4461	6144	9655
33	3604	4401	6293	9621
34	3682	4442	6221	9711
35	3611	4445	6395	9860
36	3670	4428	6409	9845
37	3770	4579	6562	9939
38	3740	4519	6646	10297
39	3780	4541	6745	11374
40	3774	4515	6879	13600

APPENDIX D

Number of Standard Errors (k) Associated With Various Confidence Levels

k	Confidence Level
1.0	.682
1.5	.866
2.0	.955
2.5	.988
3.0	.997
3.5	.999
4.0	About 1.000

APPENDIX E
Simulated Sample Results
Simple Random Sampling

Preliminary Sample - 30 Elements

60	1590	458
267	1780	5453
113	0	458
0	2006	659
0	228	413
0	10297	127
8086	3961	0
1490	82	814
593	1378	2443
620	1975	63

Mean Individual Income ($\bar{x}$) $1514
Sample Standard Deviation(s) $2452

Regular Sample - Additional 346 Elements

Total Incomes Drawn for Extended Sample	$505,708
New Mean Individual Income With 376 Indians	$1,465
New Standard Deviation With 376 Indians	$2,287
Standard Error	$118

APPENDIX F

Sample Size Required
With Simple Random Sampling
@ .95 Confidence Interval

N	A	n
1000	$50	905
	100	706
	200	376
	300	211
2000	50	1656
	100	1092
	200	462
	300	236
5000	50	3290
	100	1624
	200	537
	300	254

APPENDIX G

Simulated Sample Results
Stratified Simple Random Sampling

Strata	1	2	3	4	
	$ 0	$1289	$3373	$7833	
	409	1384	5329	9860	
	252	1135	3824	7429	
	0	2263	3284	8141	
	408	2443	4461	13600	
	13	1231	4388	8978	
	204	2309	4195	9845	
	369	1404	4321	9322	
	435	1946	4636	7216	
	267	1179	4359	11374	
	9	1932	5745	9051	
	13	1706	4321	7389	
	236	2113	4701	7525	
	188	1009	6745	9143	
	409	2888	3112	7673	
	169	1979	3086	9621	
	0	1438	4622	7745	
	368	1333	3670	9297	
	22	1022	4401	8497	
	470	1628	5329	8591	
	142	1345	3877	8197	
	358	1545	5243	9322	
	43	1109	3611	9711	
	0	1798	3800	7592	
	30	2965	5083	8086	
	282	1332	6646	8838	
	552	2817	4804	9583	
	0	1467	3222	7745	
	179	2684	6409	9939	
	93	2242	3180	8688	
Stratum Mean	$197.40	$1764.50	$4457.23	$8861.03	
Stratum Standard Deviation	$174.95	$ 587.53	$1020.76	$1333.14	
Overall Mean					$1414.73
Standard Error					$ 35.96

APPENDIX H

Simulated Sample Results
Single Stage Cluster Sampling

Clusters in Population

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50

TOTAL INCOME OF 20 INDIANS LOCATED
IN CLUSTERS SELECTED RANDOMLY

CLUSTER NUMBER DRAWN FOR SAMPLE	CLUSTER INCOME	CLUSTER TOTAL
10	$31,778	
30	31,835	
21	31,213	
19	29,840	
23	31,713	
5	29,914	
8	25,636	
7	34,057	
27	27,299	
29	37,122	
41	28,055	
6	25,317	
15	26,135	
34	30,698	
32	29,996	
39	26,182	
20	26,403	
48	26,357	
43	26,045	
Mean Income Per Cluster		$29,274
Mean Income Per Indian		$ 1,463.70
Standard Error		$ 29.43

APPENDIX I

Simulated Sample Results
Systematic Sampling

Total income of 333 elements drawn with the 2nd unit and selecting every 3rd one thereafter	$482,090
Mean individual Indian income	$1,448
Sample standard deviation (Approximate)	$2,287
Standard error	$124.97

STATE OF ARIZONA

GOVERNOR
JACK WILLIAMS

THIRTIETH STATE LEGISLATURE

President of the Senate
WILLIAM JACQUIN

Speaker of the House
TIMOTHY BARROW

DEPARTMENT OF
ECONOMIC PLANNING AND DEVELOPMENT

Board Chairman
MASON WALSH, Phoenix

Executive Director of Department
ROBERT G. WORDEN

Director of Planning Division
HARRY F. HIGGINS

Chief of Planning Section
DENNIS A. DAVIS

Chief of Economic Research Section
C. W. "JICK" MYERS